INTERNATIONAL DESIGN CONFERENCE – DESIGN 2022
https://doi.org/10.1017/pds.2022.77

Planning the Analysis of Use Phase Data in Product Planning

M. Meyer [1,✉], I. Wiederkehr [1], C. Koldewey [1] and R. Dumitrescu [1,2]

[1] Heinz Nixdorf Institute, Paderborn University, Germany, [2] Fraunhofer IEM, Germany

✉ maurice.meyer@hni.uni-paderborn.de

Abstract

The ongoing digitalization of products offers product managers new potentials to plan future product generations based on data from the use phase instead of assumptions. However, product managers often face difficulties in identifying promising opportunities for analyzing use phase data. In this paper, we propose a method for planning the analysis of use phase data in product planning. It leads product managers from the identification of promising investigation needs to the derivation of specific use cases. The application of the method is shown using the example of a manufacturing company.

Keywords: data-driven design, early design phase, product planning, big data analysis, industrial data analytics

1. Introduction

Product planning is changing. For decades, product managers have primarily relied on methods such as workshops and customer interviews to identify improvement potentials for their products (Cooper, 1988; Gausemeier *et al.*, 2019). This information was predominantly qualitative and incomplete (Bosch-Sijtsema and Bosch, 2015; Kim and Wilemon, 2002), which is why product managers had to rely on their experience and make assumptions for improving products. However, with the transformation from mechatronic products to cyber-physical systems, extensive data from the use phase are now available that product managers can examine in search of improvement potentials (Porter and Heppelmann, 2014, 2015). This enables new products to be planned based on facts instead of assumptions (Holler *et al.*, 2017; Wuest, Hribernik and Thoben, 2014; Xu, Frankwick and Ramirez, 2016).
Although this potential is widely recognized in both research and practice, e.g., see Dumitrescu *et al.* (2021), there is a lack of suitable approaches for implementation in practice, especially for manufacturing companies. For illustration, Bertoni points out that most publications on the analysis of use phase data address data from social media and online reviews. Sensor data, which are prevalent in the manufacturing industry and its CPS, are only considered in a minor number of articles (Bertoni, 2020). For product managers in the manufacturing industry, planning promising data analyses is already a major challenge. Meyer *et al.* show that they have considerable difficulties in developing promising use cases for the analysis of data from the use phase. They need methodological support for planning data analyses (Meyer and Panzner *et al.*, 2021). Therefore, in this paper, we address the research question of how to successfully plan the analysis of use phase data in product planning.
Our research design is based on the design science guidelines proposed by Hevner, March and Park (2004). They describe design science as a problem-solving process that aims at creating an innovative, purposeful artifact for a specified and relevant problem domain. Through this process, knowledge and understanding of both the problem and its solution are acquired (Hevner, March and Park, 2004). In our paper, we present a method (artifact) for planning the analysis of use phase data in product planning (problem domain).

The paper is structured as follows: In section 2, the scientific background of analyzing use phase data in product planning is described. The proposed method is introduced in section 3. In section 4, its application with a manufacturing company is shown. Section 5 focuses on the validation of the proposed method. Section 6 discusses the contribution of the method and future research needs.

2. Scientific Background

In this section, the scientific background of analyzing use phase data in product planning is described. First, its three basic concepts *product planning*, *use phase data*, and *data analytics* are addressed. Second, the purpose of analyzing use phase data in product planning is discussed. Third, a reference process model is presented that includes the planning of use phase data analyses in its first main process.

2.1. Basic Concepts

Analyzing use phase data in product planning builds on three basic concepts: *product planning*, *use phase data*, and *data analytics*. In the following, these are briefly described.

2.1.1. Product Planning

Product planning represents the first phase in the product creation process. The main tasks are the identification of the success potentials of the future, the discovery of promising product ideas as well as business planning (Gausemeier *et al.*, 2011; Koen *et al.*, 2001). Results are the products to be developed by the company (Ulrich and Eppinger, 2016) and the corresponding requirements lists (Pahl *et al.*, 2007). Some authors refer to product planning as the initial phase or phase zero of product development (Ulrich and Eppinger, 2016). Other authors like Koen et al. call it the front end of innovation or the fuzzy front end (Koen *et al.*, 2001). Cooper and Kleinschmidt show that product planning activities determine the success of a new product (Cooper and Kleinschmidt, 1986). In a recent publication, Cooper counts them among the drivers of success in new-product development (Cooper, 2019).

2.1.2. Use Phase Data

Every product goes through a typical lifecycle. In product lifecycle management, the product lifecycle categorizes the following three major phases: beginning of life (BOL), middle of life (MOL), and end of life (EOL). The MOL includes the minor phases *use*, *service*, and *maintenance* (Kiritsis, 2011; Terzi *et al.*, 2010).

For the data generated and collected in the use phase, different classifications exist in the literature. Beverungen et al. divide these data into usage data, context data, and status data (Beverungen *et al.*, 2019). Kreutzer calls the data generated in the use phase field data and divides it into three classes: Technical measurements, user data, and system data (Kreutzer, 2019). Balasubramanian et al. divide the use phase data of cars into five types: External conditions, technical status of equipment, product usage, personal data and preferences, and direct communications (Balasubramanian *et al.*, 2016). Hou and Jiao distinguish user-generated data, product operating data, and environmental data (Hou and Jiao, 2020). A comparison of the classifications shows that there are overlaps and similarities of classes, e.g., the division into user-oriented, product-oriented, and context-oriented data. However, due to its uniqueness, each classification makes an important contribution to the understanding of the term *use phase data*.

2.1.3. Data Analytics

Technological advancements in big data architectures as well as artificial intelligence and machine learning enable the efficient analysis of big data (Lueth *et al.*, 2016). The process of accessing, aggregating, and analyzing large amounts of data from multiple sources is called data analytics (DA). DA enables companies to extract knowledge from data to understand historical and predict future events (Tyagi, 2003). The foundations of DA are mathematics, computer science, and business analysis techniques (Porter and Heppelmann, 2015). In addition, Reinhart stresses the importance of integrating specific domain knowledge, e.g., about manufacturing (Reinhart, 2016). DA can be divided into four types with increasing value and complexity: descriptive, diagnostic, predictive, and prescriptive (Steenstrup *et al.*, 2014). LaValle *et al.* surveyed the challenges and opportunities of data analytics with more than 3000 business executives, managers, and analysts (LaValle *et al.*, 2011). Among others, they found that top-performing companies utilize DA five times more than lower-performing companies. The use of DA ranges from management tasks to day-to-day operations; however, the authors stress the correlation found for analytics-driven management and business performance (LaValle *et al.*, 2011).

2.2. Analyzing use Phase Data in Product Planning

Product planning aims at a sharp definition of the product to be developed (Cooper, 2019). However, Herstatt and Verworn among others describe that product planning and its decisions are characterized by great uncertainties (Herstatt and Verworn, 2004). This results in a product definition based on assumptions instead of facts. At the same time, decisions made during product planning have a considerable influence on all later phases, e.g., by determining their costs. After product planning, changes to the product require increasingly more effort and costs. Therefore, it is necessary to dissolve as many uncertainties as possible within product planning to create a sharp product definition that does not require major changes in later lifecycle stages (Herstatt and Verworn, 2004).

The uncertainties of product planning do not arise anew with every planned product. Rather, they accumulate over generations. In a study with 247 engineers, Albers et al. found that only 7% of the developments concern entirely new products. In contrast, 93% of the developments address the further development of existing products and thus the development of a new product generation (Albers *et al.*, 2014). The assumptions resulting from the uncertainties of product planning can thus be carried along over several product generations. For this reason, products must be reviewed regarding both old and new uncertainties and assumptions.

Information and data from the use phase are suitable for clarifying some of the uncertainties and assumptions. Traditionally, the acquisition of high-quality use phase information has been difficult (Deng *et al.*, 2021); however, due to the transformation of mechatronic products to cyber-physical systems, extensive use phase data from the product generations in the field are now available (Porter and Heppelmann, 2014). These data can be investigated with statistical analysis, data mining, and machine learning methods (Hou and Jiao, 2020; Igba *et al.*, 2015). The results of the data analysis lead to new insights about the product and its users (Meyer and Wiederkehr *et al.*, 2021). For the planning of a future, improved product generation, they enable decisions based on facts instead of assumptions (Holler *et al.*, 2017; Wuest, Hribernik and Thoben, 2014; Xu, Frankwick and Ramirez, 2016). Figure 1 illustrates these concepts.

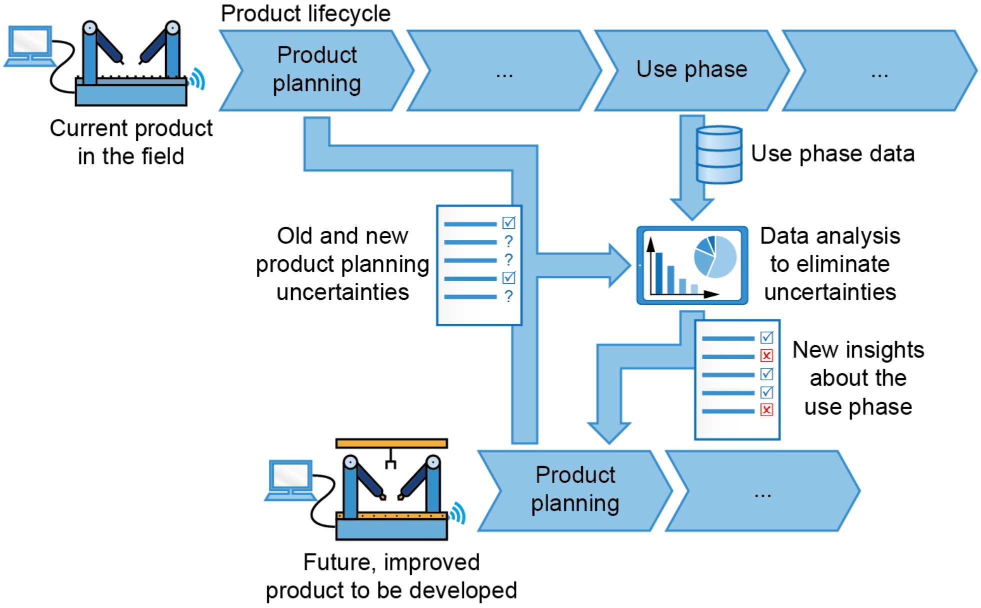

Figure 1. Analysis of use phase data to eliminate uncertainties in product planning

https://doi.org/10.1017/pds.2022.77 Published online by Cambridge University Press

2.3. Reference Process Model

Meyer *et al.* (2022) propose a reference process model for the analysis of use phase data in product planning. It consists of four main processes: (1) Planning of the use phase data analysis, (2) Analysis and data preparation, (3) Analytics workflow design and modeling, and (4) Exploitation of the data analysis results (Meyer *et al.*, 2022).

(1) The planning of the use phase data analysis aims at promising use cases. For this purpose, investigation needs are identified and the data analytics capabilities are analyzed. Subsequently, the objectives and boundary conditions of the investigation are defined and concrete questions for the data analysis are developed. (2) The second main process addresses the preparation of the data analysis. Therefore, the previously defined questions are translated into specific objectives for the data scientist. Subsequently, the required data are defined, collected, and described. (3) The goal of the third main process is the results of the data analysis. For this purpose, data analytics workflows are designed, the data are pre-processed, a model is built, and finally validated. (4) In the last main process, the results of the data analysis are exploited in product planning. First, the results are interpreted and transformed into new insights about the product. Then, ideas for improving the product and related requirements are developed. Finally, the product strategy and other documents are updated based on the new insights, ideas, and requirements (Meyer *et al.*, 2022). Figure 2 shows the reference process model.

For product managers who want to analyze use phase data to improve their products, even the *planning* main process presents a challenge. In another paper, Meyer et al. describe an experiment in which product managers are asked to develop use cases for the analysis of data from the use phase in different settings. In the process, the product managers involved are confronted with numerous difficulties, such as identifying promising questions. Based on these findings, the authors conclude that product managers need methodological support for planning the analysis of use phase data (Meyer and Panzner *et al.*, 2021). To the best of our knowledge, no method addresses all sub-processes of the reference process model's *planning* main process. Therefore, we propose a new method for planning use phase data analyses in product planning in this paper.

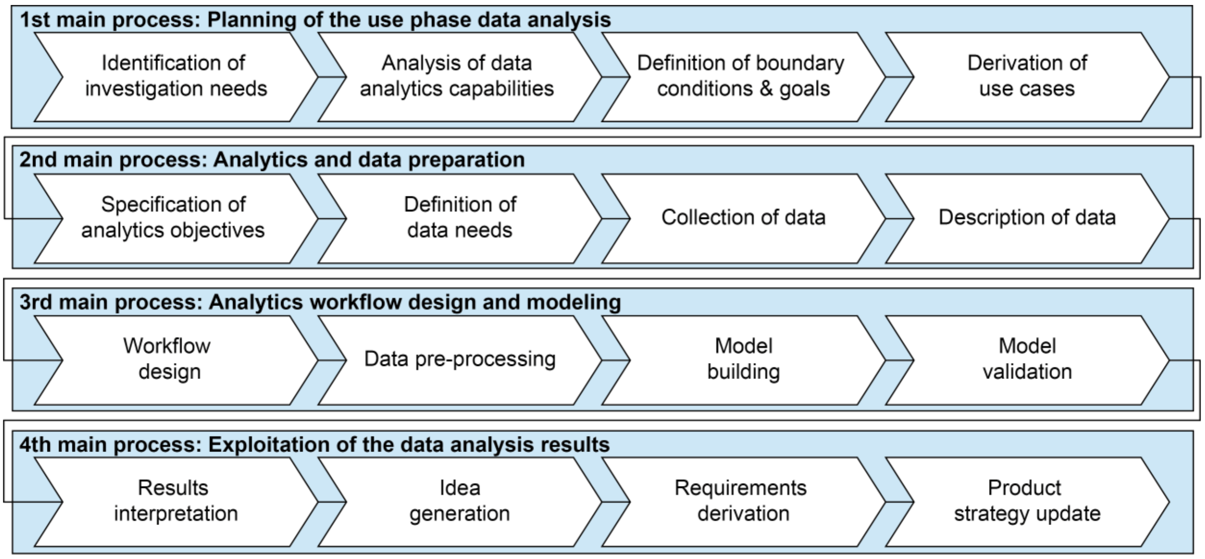

Figure 2. Reference process model for analyzing use phase data in product planning based on Meyer *et al.* (2022) (updated)

3. Method

The method consists of four phases that correspond to the four sub-processes of the *planning* main process described above. Figure 3 shows the corresponding phases, tasks, and results. In the following, the proposed tasks for each phase are briefly described.

https://doi.org/10.1017/pds.2022.77 Published online by Cambridge University Press

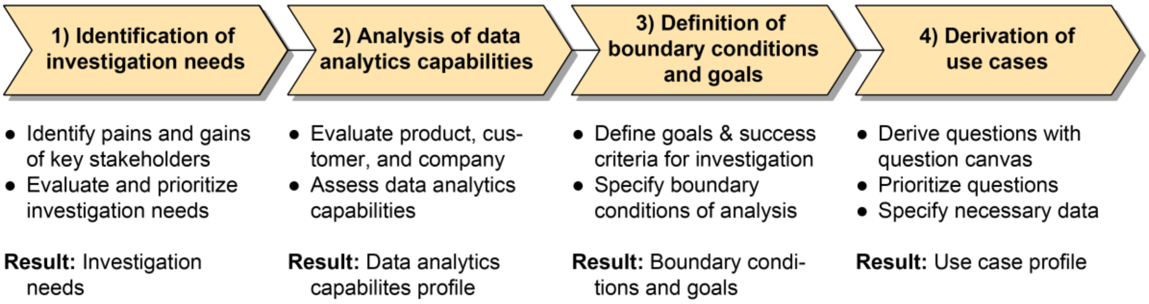

Figure 3. Phases, tasks, and results of planning use phase data analyses in product planning

1) Identification of investigation needs: The first phase involves identifying investigation needs that are characterized by both high strategic relevance and high uncertainty. For the product to be examined, first, the customer profile of the Value Proposition Canvas according to Osterwalder *et al.* (2014) is applied: For key stakeholders (e.g., customers, users, service), the pains and gains associated with the product are identified. These represent natural investigation needs that are then evaluated and prioritized in terms of their strategic relevance and their uncertainty.

2) Analysis of data analytics capabilities: In the second phase, the data analytics capabilities for the company and the considered product are evaluated. For this, seven criteria in three areas apply: In the product area, the determinacy of product behavior, the availability of intelligent components, and the comparability of product instances in the field (Holler *et al.*, 2016). The customer area addresses the available options for data access as well as the willingness of customers to share data (Holler *et al.*, 2016; Wilberg *et al.*, 2017). The company area evaluates the data consistency as well as the data competencies available in the company (Wilberg *et al.*, 2017). All seven criteria are first answered textually. Then, each capability is assessed on a 4-level scale from very low to very high. The resulting profile highlights weaknesses and strengths to be considered when planning the analysis.

3) Definition of boundary conditions and goals: The third phase serves to define the boundary conditions and goals. It builds on the results of the first two phases. For each selected investigation need, the goal of the investigation and success criteria are first defined. Then, building on the capabilities profile, the boundary conditions of the analysis are described. For that, four guiding questions are asked: (1) What are the requirements for the data analysis? (2) What assumptions can be made so that they can be initially disregarded in the analysis? (3) What system boundaries shall be set for the analysis? (4) What restrictions and risks are to be expected? The questions build on the suggestions of Chapman *et al.* (2000) and Verein Deutscher Ingenieure e.V. (2019). The answers to these questions are essential information for data scientists. Also, they are indispensable for interpreting the data analysis results.

4) Derivation of use cases: The fourth phase addresses the derivation of use cases for the analysis. For each planned investigation, specific questions are derived. In conjunction, the goal, the boundary conditions, and the questions represent a use case for analyzing use phase data in product planning. For deriving promising questions, a *question canvas* is proposed. This canvas visually connects the known (facts) with the unknown (questions) by asking the questions "What is known? What do we know?" in the center and the questions "What do we not know? What do we want to find out?" outside of the center. First, all information that is known about the use case is gathered in the middle (e.g., "The errors only occur at night"). Based on this, questions are formulated that build directly on known information (e.g., "What materials does the machine process at night?") or address gaps in knowledge (e.g., "At what ambient temperatures do the defects occur?"). The questions are then prioritized. For the selected questions, the necessary data to answer them are written down. The result of this phase and the overall method is a use case profile that provides the necessary information for the data scientist to prepare the data analysis (see 2nd main process of the reference process model displayed in Figure 2).

4. Exemplary Application

So far, we have performed the method with four manufacturing companies. In this section, we describe its application with a manufacturing company that offers systems for banking and commerce. In this example, we considered an ATM product family. The method was performed in a workshop.

https://doi.org/10.1017/pds.2022.77 Published online by Cambridge University Press

1) Identification of investigation needs: At the beginning of the workshop, the investigation needs were identified. For this purpose, the pains and gains of the three stakeholders *customers* (banks), *users* (bank customers), and *sales employees* were recorded using brainstorming. Then, the workshop participants evaluated all pains and gains in a group discussion concerning their strategic relevance and their associated uncertainty. We used a portfolio to illustrate this evaluation and help the prioritization. The portfolio distinguishes low, medium, and high priority investigation needs. Investigation needs characterized by high levels of strategic relevance and uncertainty have the highest priority. They are located in the upper right corner of the portfolio and must be primarily investigated with data analyses. In the workshop, for example, the reliability of the considered ATM was assessed as highly relevant. Also, the workshop participants assumed a medium level of uncertainty. Among others, this specific investigation need was selected for further elaboration. Figure 4 shows the described portfolio for prioritizing the identified investigation needs.

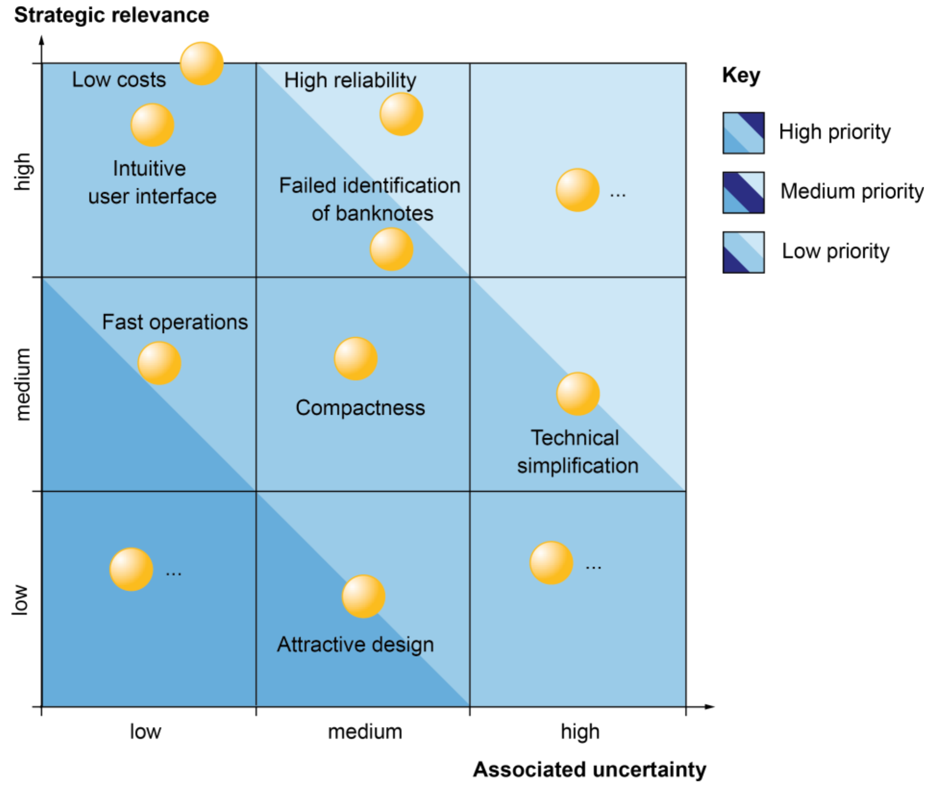

Figure 4. Portfolio for prioritizing the identified investigation needs

2) Analysis of data analytics capabilities: The analysis of the analytics capabilities showed that the company under consideration has very good prerequisites for analyzing use phase data. Only the determinacy of the system's behavior was rated as very low. This is mainly due to the environmental conditions: For example, the banknotes to be withdrawn and dispensed may be made of relatively strong and thick paper in one country and very thin paper in another. In addition, customers often put items into the machine that are not banknotes. Such outside influences often make ATM behavior unpredictable and thus must be considered to obtain valid results. In contrast, data competencies received the best rating (very high). The company under review has a data science department. Many other departments have also already made connections with data analyses so that both analytics experience and a certain affinity for the analysis of data are present. All other criteria were rated as high.

3) Definition of boundary conditions and goals: Subsequently, the workshop participants discussed the goal and boundary conditions for analyzing the reliability of the product under consideration. They agreed on the goal to identify the most serious causes of failure as these directly affect reliability. Among the criteria defined to confirm the success of the analysis is that the existing assumptions are confirmed or refuted by analysis. For the definition of the boundary conditions, we followed the four guiding questions. (1) An important requirement for the analysis is that environmental factors must be

https://doi.org/10.1017/pds.2022.77 Published online by Cambridge University Press

considered because they have a direct influence on the behavior and comparability of the product instances (as discussed above). (2) Initial assumptions are, for example, that all product instances considered were installed correctly and had no transport damage. If the results of the analysis could not be explained, these assumptions would have to be reviewed. (3) As system boundaries, workshop participants determined that only product instances of the latest generation should be considered as older instances would not be comparable. (4) To conclude this phase, the participants collected the restrictions and risks of the investigation. For example, they again emphasized that environmental conditions can sometimes fluctuate significantly and thus influence analysis results. The results of this phase are displayed in the use case profile in Figure 5.

4) Derivation of use cases: Finally, the workshop participants expanded the selected investigation need into a concrete use case. For this, they first collected their prior knowledge about the investigation need in the question canvas, e.g., that service technicians often report that reels break in the product. Then, they collected questions about these known points as part of a question storming session. For example, workshop participants wanted to find out what processes precede the reel breaks, whether there were any anomalies in sensor readings beforehand, and whether some environmental conditions promote earlier reel breaks. Answers to these questions would help the workshop participants isolate the problem and identify the causes of the failures. Finally, the workshop participants recorded the necessary data sets to answer the questions. After adding the questions and data, the use case was fully described and ready to be sent to the data scientists. Figure 5 shows the questions and data in the use case profile.

<table>
<tr><th colspan="6">Use Case: Failure analysis for improving ATM reliability</th></tr>
<tr><th colspan="2">Goal</th><th colspan="2">Requirements</th><th colspan="2">Assumptions</th></tr>
<tr><td colspan="2">Identification of the most serious causes of failure</td><td colspan="2">▪ Consideration of environmental factors
▪ No individual cases, but solid and representative analysis basis
▪ …</td><td colspan="2">▪ Correct installation of all considered product instances
▪ No transport damages
▪ …</td></tr>
<tr><th colspan="2">Success criteria</th><th colspan="2">System boundaries</th><th colspan="2">Restrictions and risks</th></tr>
<tr><td colspan="2">▪ Solid numerical basis for the problem (>100 ATMs)
▪ Confirmed or refuted assumptions
▪ …</td><td colspan="2">▪ Only product instances of the latest generation
▪ No consideration of products without equivalent trace data
▪ …</td><td colspan="2">▪ Fluctuation of environmental conditions
▪ Missing necessary data points
▪ …</td></tr>
<tr><th colspan="3">Questions</th><th colspan="3">Necessary data</th></tr>
<tr><td colspan="3">After how many runtime hours do the reels break?</td><td colspan="3">EPROM, service reports, ticket system data</td></tr>
<tr><td colspan="3">Are there anomalies in sensor measurements before the reel break?</td><td colspan="3">Sensor und actuator data</td></tr>
<tr><td colspan="3">Do certain environmental conditions promote an early reel break?</td><td colspan="3">Sensor und actuator data</td></tr>
<tr><td colspan="3">…</td><td colspan="3">…</td></tr>
</table>

Figure 5. Use case profile as the final result of the method

5. Validation

The validation of the method is performed using the validation square by Pedersen *et al.* (2000). It describes four criteria for the validity of design methods: (1) theoretical structural validity, (2) empirical structural validity, (3) empirical performance validity, and (4) theoretical performance validity. In the following, the criteria are described and applied to evaluate the validity of the presented method.

(1) Theoretical structural validity describes the acceptability of the individual elements of the method as well as the method as a whole. The first can be demonstrated via the literature and citations. The second can be examined via an analysis of the information flows between the elements of the method (Pedersen *et al.*, 2000). **Evaluation**: The phases of the method correspond to the process for planning use phase data analyses according to Meyer *et al.* (2022). Within the phases of the method,

https://doi.org/10.1017/pds.2022.77 Published online by Cambridge University Press

well-known and often cited papers of other authors are taken up, e.g., the Value Proposition Canvas (Osterwalder *et al.*, 2014). The individual tasks and phases as well as their results are synchronized and build on each other. This can be seen in Figure 3, for example. For these reasons, we consider both the individual elements of the method and the method as a whole to be theoretically structurally valid.

(2) Empirical structural validity exists if the selected application examples are suitable for the validation of the method. Here it needs to be shown that the application examples and their problems represent the problem class for which the method was designed (Pedersen *et al.*, 2000). **Evaluation**: So far, we have applied the presented method with four manufacturing companies. All of them pursue the goal of identifying improvement potentials for future product generations with the help of use phase data analyses. In the beginning, they did not know what exactly they should analyze, i.e. which questions they should pursue. This initial situation corresponds to the problem class that we address with the method. Therefore, we consider the empirical structural validity to be fulfilled.

(3) Empirical performance validity indicates that the method fulfills its purpose and leads to the desired results for the considered application examples. The added value can also be quantified by a performance comparison with alternative methods for the considered problem (Pedersen *et al.*, 2000). **Evaluation**: For more than three years, we have accompanied the four companies with which we have applied the method in a research project. At the start of the project, the identification of promising use cases for the analysis of use phase data was a central challenge. Before developing this method, we had already designed a hypothesis-based approach (Meyer *et al.*, 2020), which, however, had some drawbacks in practical application, e.g., that a large proportion of the hypotheses posed were not testable and that the quantification of the hypotheses was often very unreliable. By applying the method proposed in this paper, all four companies were able to identify promising use cases and proceed to the next steps (i.e., data collection and analysis). Since all of the four companies are now acquiring and analyzing data, we consider the empirical performance validity to be given, even if we cannot quantify the superiority of the presented method.

(4) Theoretical performance validity concerns the generalizability of the performance of the method to further cases beyond the application examples. Pedersen *et al.* (2000) suggest investigating this criterion by inductive reasoning based on the other three criteria to test the confidence to general usefulness. **Evaluation**: The four companies with which we have already used the method have very different characteristics (e.g., company size, experience with data analytics, number of products in the field, amount of data per product). Nevertheless, the method has helped all companies to develop promising use cases, which are now being successfully pursued. Therefore, we assume that the presented method would also be effective at further manufacturing companies, and we are confident that the method achieves theoretical performance validity.

6. Contribution and Future Research Needs

In this paper, we propose a method for planning use phase data analyses in product planning. It guides users to promising use cases that address specific product planning uncertainties. At the same time, it ensures that the use cases are realistic and feasible by analyzing the data analytics capabilities from seven perspectives. Using the method, product managers are enabled to plan data analyses that lead to valuable insights about the considered product. Using the validation square by Pedersen et al. (2000), we demonstrated the validity of the method. Thus, for practice, this paper provides a validated method for the planning of use phase data analyses in product planning. As four manufacturing companies have already successfully performed the method, practical suitability has been shown. For research, we propose an innovative, purposeful, and tested artifact for a specified and relevant problem domain and thereby meet the demands of design science research (Hevner, March and Park, 2004).

With this in mind, we see three main future research needs: First, as the generalizability of the method beyond the four application examples can only be assumed in this paper, the method should be applied with further manufacturing companies. Other contexts could bring up new requirements for the method that are not yet fulfilled. Second, as soon as the first companies have fully conducted the reference process for analyzing use phase data in product planning (see Figure 2), a review of the planning phase should be conducted. This should include, for example, an assessment of whether important aspects for

https://doi.org/10.1017/pds.2022.77 Published online by Cambridge University Press

later main processes were not sufficiently addressed in the planning main process. Third, specific examples of analyzing use phase data in product planning should be developed which demonstrate possible applications and the associated added value. Such application examples can serve as orientation for users of the method. Simultaneously, they show the potential of such analyses. The latter is important in inexperienced companies that are unsure whether the potentials are worth the efforts.

Acknowledgments

This publication is funded by the German Federal Ministry of Education and Research (BMBF). The authors are responsible for the content of this publication.

References

Albers, A., Bursac, N., Urbanec, J., Lüdcke, R. and Rachenkova, G. (2014), "Knowledge Management in Product Generation Development – an empirical study", *Design for X – Beiträge zum 25. DfX-Symposium*, TUHH Universitätsbibliothek, Hamburg, Germany, pp. 13–24.

Balasubramanian, J., Beiker, S., Chauhan, S., Colombo, T., Hansson, F. et al. (2016), Car data: paving the way to value-creating mobility: Perspectives on a new automotive business model, McKinsey&Company.

Bertoni, A. (2020), "Data-Driven Design in Concept Development: Systematic Review and Missed Opportunities", *Proceedings of the Design Society: DESIGN Conference*, Cavtat, Croatia, Cambridge University Press, pp. 101–110. https://dx.doi.org/10.1017/dsd.2020.4

Beverungen, D., Müller, O., Matzner, M., Mendling, J. and vom Brocke, Jan (2019), "Conceptualizing smart service systems", *Electronic Markets*, Vol. 29 No. 1, pp. 7–18. https://dx.doi.org/10.1007/s12525-017-0270-5

Bosch-Sijtsema, P. and Bosch, J. (2015), "User Involvement throughout the Innovation Process in High-Tech Industries", *Journal of Product Innovation Management*, Vol. 32 No. 5, pp. 793–807.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T. et al. (2000), "CRISP-DM 1.0 – Step-by-step data mining guide", SPSS, Vol. 9.

Cooper, R.G. (1988), "Predevelopment activities determine new product success", *Industrial Marketing Management*, Vol. 17 No. 3, pp. 237–247. https://dx.doi.org/10.1016/0019-8501(88)90007-7

Cooper, R.G. (2019), "The drivers of success in new-product development", *Industrial Marketing Management*, Vol. 76, pp. 36–47. https://dx.doi.org/10.1016/j.indmarman.2018.07.005

Cooper, R.G. and Kleinschmidt, E.J. (1986), "An Investigation into the New Product Process: Steps, Deficiencies, and Impact", *Journal of Product Innovation Management*, Vol. 3 No. 2, pp. 71–85.

Deng, Q., Wellsandt, S., Hribernik, K. and Thoben, K.-D. (2021), "Understanding Users and Products in Product Development: The Application of Product Usage Information and its Challenges", *Proceedings of the International Conference on Engineering Design (ICED21)*, Gothenburg, Sweden, pp. 3299–3308.

Dumitrescu, R. et al. (2021) Engineering in Germany – The status quo in business and science, a contribution to Advanced Systems Engineering. Paderborn, Germany.

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.

Gausemeier, J., Dumitrescu, R., Echterfeld, J., Pfänder, T., Steffen, D. et al. (2019), *Innovationen für die Märkte von morgen: Strategische Planung von Produkten, Dienstleistungen und Geschäftsmodellen*, Carl Hanser Verlag, München.

Herstatt, C. and Verworn, B. (2004), "The 'Fuzzy Front End' of Innovation", *Bringing Technology and Innovation into the Boardroom*, Palgrave Macmillan, London, pp. 347–372. https://dx.doi.org/10.1057/9780230512771

Hevner, A., March, S.T. and Park, J. (2004), "Design Science in Information Systems Research", *MIS Quarterly*, Vol. 28 No. 1, pp. 75–105.

Holler, M., Stoeckli, E., Uebernickel, F. and Brenner, W. (2016), "Towards Understanding closed-loop PLM: The Role of Product Usage Data for Product Development enabled by intelligent Properties", *BLED 2016 Proceedings*, Bled, Slovenia, pp. 479–491.

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", *Proceedings of the 50th Hawaii International Conference on System Sciences (HICSS)*, Hawaii, USA, pp. 4726–4735.

Hou, L. and Jiao, R.J. (2020), "Data-informed inverse design by product usage information: a review, framework and outlook", *Journal of Intelligent Manufacturing*, No. 31 Vol. 3, pp. 529–552.

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://dx.doi.org/10.1016/j.rcim.2014.12.004

Kim, J. and Wilemon, D. (2002), "Focusing the fuzzy front-end in new product development", *R&D Management*, Vol. 32 No. 4, pp. 269–279. https://dx.doi.org/10.1111/1467-9310.00259
Kiritsis, D. (2011), "Closed-loop PLM for intelligent products in the era of the Internet of things", *Computer-Aided Design*, Vol. 43 No. 5, pp. 479–501. https://dx.doi.org/10.1016/j.cad.2010.03.002
Koen, P., Ajamian, G., Burkart, R., Clamen, A., Davidson, J. et al. (2001), "Providing Clarity and A Common Language to the "Fuzzy Front End"", *Research-Technology Management*, Vol. 44 No. 2, pp. 46–55.
Kreutzer, R. (2019), *Methodik zur Bestimmung der Nutzenpotenziale von Felddaten cyber-physischer Systeme*, [PhD Thesis], RWTH Aachen.
LaValle, S., Lesser, E., Shockley, R., Hopkins, M.S. and Kruschwitz, N. (2011), "Big Data, Analytics and the Path From Insights to Value", *MIT Sloan Management Review*, Vol. 52 No. 2, pp. 21–31.
Lueth, K.L., Patsioura, C., Williams, D.Z. and Kermani, Z.Z. (2016), Industrial Analytics 2016/2017: The current state of data analytics usage in industrial companies, Digital Analytics Association e.V. Germany (DAAG), IoT Analytics GmbH.
Meyer, M., Frank, M., Massmann, M., Wendt, N. and Dumitrescu, R. (2020), "Data-Driven Product Generation and Retrofit Planning", *Procedia CIRP, Vol. 93*, Chicago, Illinois, USA, pp. 965–970. https://dx.doi.org/10.1016/j.procir.2020.04.122
Meyer, M., Panzner, M., Koldewey, C. and Dumitrescu, R. (2021), "Towards Identifying Data Analytics Use Cases in Product Planning", *Procedia CIRP, Vol. 104*, Athens, Greece, pp. 1179–1184. https://dx.doi.org/10.1016/j.procir.2021.11.198
Meyer, M., Wiederkehr, I., Koldewey, C. and Dumitrescu, R. (2021), "Understanding Usage Data-Driven Product Planning: A Systematic Literature Review", *Proceedings of the International Conference on Engineering Design (ICED21)*, Gothenburg, Sweden, pp. 3289–3298. https://dx.doi.org/10.1017/pds.2021.590
Meyer, M., Wiederkehr, I., Panzner, M., Koldewey, C. and Dumitrescu, R. (2022), "A Reference Process Model for Usage Data-Driven Product Planning", *Proceedings of the 55th Hawaii International Conference on System Sciences (HICSS)*, Hawaii, USA, 6105-6114.
Osterwalder, A., Pigneur, Y., Bernarda, G., Smith, A. and Papadakos, P. (2014) *Value proposition design: How to create products and services customers*. Wiley, Hoboken, NJ, USA.
Pahl, G., Beitz, W., Feldhusen, J. and Grote, K.-H. (2007), *Engineering design: A systematic approach*, 3rd edn., Springer, London. https://dx.doi.org/10.1007/978-1-84628-319-2
Pedersen, K., Emblemsvåg, J., Bailey, R., Allen, J.K. and Mistree, F. (2000), "Validating Design Methods and Research: The Validation Square", *Proceedings of DETC '00*, 2000 *ASME Design Engineering Technical Conferences*, American Society of Mechanical Engineers, New York, NY, pp. 379–390.
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. 96–114.
Reinhart, F. (2016), "Industrial Data Science - Data Science in der industriellen Anwendung", *Industrie 4.0 Management*, Vol. 32 No. 6, pp. 27–30.
Steenstrup, K., Sallam, R., Eriksen, L. and Jacobson, S. (2014), Industrial Analytics Revolutionizes Big Data in the Digital Business, *Gartner*, Available at: https://www.gartner.com/en/documents/2826118/industrial-analytics-revolutionizes-big-data-in-the-digi (Accessed: 21 November 2020).
Terzi, S., Bouras, A., Dutta, D., Garetti, M. and Kiritsis, D. (2010), "Product lifecycle management – from its history to its new role", *International Journal of Product Lifecycle Management*, Vol. 4 No. 4, pp. 360–389.
Tyagi, S. (2003), "Using data analytics for greater profits", *Journal of Business Strategy*, Vol. 24 No. 3, pp. 12–14. https://dx.doi.org/10.1108/02756660310734938
Ulrich, K.T. and Eppinger, S.D. (2016), *Product design and development*, 6th edn., McGraw-Hill, New York, NY.
Verein Deutscher Ingenieure e.V., VDI/VDE 3714 Blatt 1 (2019): Implementation and operation of big data applications in the manufacturing industry – Implementation of big data projects, Verein Deutscher Ingenieure e. V., Beuth Verlag, Berlin.
Wilberg, J., Triep, I., Hollauer, C. and Omer, M. (2017), "Big Data in product development: Need for a data strategy", *Proceedings of PICMET '17: Technology Management for Interconnected World*, Portland State University, Portland, USA, pp. 1–10.
Wuest, T., Hribernik, K. and Thoben, K.-D. (2014), "Capturing, Managing and Sharing Product Information along the Lifecycle for Design Improvement", *Proceedings of the 10th International Workshop on Integrated Design Engineering*, Magdeburg/Gommern, 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://dx.doi.org/10.1016/j.jbusres.2015.10.017

https://doi.org/10.1017/pds.2022.77 Published online by Cambridge University Press