Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 128 (2024) 114–119

www.elsevier.com/locate/procedia

34th CIRP Design Conference

Methodical Support for Identifying and Selecting Data and Analysis Tools in the Data-Driven Development of Complex Mechatronic Systems

Steffen Wagenmann[a*], Jakob Rall[b], Artur Krause[c], Felicia Weidinger[a], Lutz Sommer[b], Nikola Bursac[c], Albert Albers[a]

[a]*IPEK, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany*
[b]*Albstadt University of Applied Sciences Ravensburg-Weingarten, 88250 Weingarten, Germany*
[c]*ISEM, Hamburg University of Technology, 21073 Hamburg, Germany*

*Corresponding author. Tel.: +49-7156-303-355667. *E-mail address:* steffen.wagenmann@trumpf.com

Abstract

Current research shows that analyzing usage data from reference systems holds a high potential for product development. The process model for data-driven validation of systems of objectives supports developers in performing such data analyses. However, developers still struggle in the process of identification and selection of data and analysis tools. Thus, this work aims to develop methodical support in this process. nine expert interviews are conducted from which 15 influencing factors on the process are identified. Based on this methodical support is derived. For successful process execution, three areas of expertise are identified that must be covered regarding development domain, data, and conducting analysis. Developers have differing skills and thus require distinct levels of assistance in different process stages. To guarantee support to as many developers as possible while ensuring applicability and efficiency, five personas are derived covering frequently occurring expertise patterns. Furthermore, metadata documentation is identified as a central lever to reduce complexity in handling data, which in turn decreases dependency on experts. Thus, a consumer-focused data catalog concept is developed. Activity profiles are used to guide developers through the process while providing support at different stages as needed. Finally, the evaluation of the solution by conducting workshops with 15 developers shows an improvement of 41.7 percent in task accomplishment.

© 2024 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of the 34th CIRP Design Conference

Keywords: Data Analytics¸Methodical Support; Design methodology, tools and technologies

1. Introduction

In today's manufacturing industry, increasing amounts of data are generated and collected. Analyzing customer usage data can assist in identifying requirements for new product generations, reducing uncertainty and risks in the development process. [1]. However, merely storing data can result in a "data graveyard", where substantial amounts of data remain unused. This results in high costs and consumes valuable resources. Profiting from usage data requires data analyses to generate valuable insights regarding customer behavior which in turn can be used in the development of new product generations. Data scientists face challenges in interpreting analysis results for complex mechatronic systems due to insufficient understanding of the technical system. Thus, according to Wagenmann et. al., the analysis should be carried out by an expert of the technical system itself. Therefore, it is essential to enable developers to actively incorporate data into their daily decision-making processes [3]. In the validation of requirements, approximately 67 percent of the total time spent on data analysis is allocated to interdisciplinary collaboration and the tasks of identifying and selecting usable data and analysis tools [4]. This can be traced back to a huge amount and variety of data and analysis tools. Especially machine data features complex characteristics, regarding format, source, and volume [2]. Furthermore, developers are hindered by a lack of

2212-8271 © 2024 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of the 34th CIRP Design Conference
10.1016/j.procir.2024.04.002

overview, understanding, and accessibility of existing data and analysis tools [4]. This suggests that the inefficiency in performing data analysis can be explained by a lack of methodical support for developers for the identification and selection of data and analysis tools.

2. Theoretical Background

2.1. Product Generation Engineering

Product development is seen as a complex system in which inputs are transformed into outputs through an operation system [6]. The system of objectives contains all goals, interactions, constraints, and requirements to develop a system of objects. During the development process, the system of objectives is extended and refined based on new information [5]. Thereby, continuous validation of the intermediate results against the system of objectives is the key to successful product development and risk minimization [6,7,8]. The model of product generation engineering (PGE) states that today's product development projects are always based on at least one already existing technical system, referred to as a reference system [9]. Adopting a higher proportion of shape and solution principles from the reference system increases development efficiency and reduces associated risks [10, 11]. Reference system elements often generate usage data which is collected and stored enabling developers to validate requirements defined by the system of objectives of new product generations.

2.2. Data-Driven Decision Making

According to Miller [12], Data-Driven Decision Making refers to "a decision-making process which involves collecting data, extracting patterns and facts from that data, and utilizing those facts to make inferences that influence decision-making." Several studies show that most organizational decisions are still based on intuition and experience rather than analyzed data [13, 14]. To extract value from data, it must be analyzed to generate knowledge. The availability of suitable tools for data processing is hereby crucial to condense large volumes of data [14, 15]. In complex mechatronic system development, analysis use cases can range from simple reports to predictive modeling and optimization algorithms based on advanced statistical models. Once the analysis is complete, the analysis results must be interpreted in the domain context by the decision-maker [16]. In the case of complex mechatronic systems, this requires extensive technical expertise [14, 17]. A study by MIT Sloan Review, involving 3,000 participants, revealed that obstacles to using data in decision-making were predominantly managerial and cultural, not data- or technology-related. Key barriers identified included a lack of understanding of how analytics could enhance business, insufficient staff skills for data analysis, and inadequate knowledge sharing within the organization [18, 19]. Therefore, fostering a data-sharing culture within the organization is crucial [14].

2.3. Process Model for the Data-Driven Validation of the System of Objectives

In the data-driven development of complex mechatronic systems, the combination of knowledge from different domains plays a central role for success. In particular, three different domains can be highlighted: knowledge about the data, the execution of analyses, and the product development domain itself [4]. These three domains are rarely found in a single individual, necessitating intensive interdisciplinary collaboration. Data scientists frequently face challenges in accurately interpreting data and analysis results within the technical domain, while experts in technical systems often lack data analysis skills [3]. As a result, most of the collected data remain unused in organizations [20].

To leverage the potentials of data-driven decision-making, Wagenmann et al. proposed a process model for data-driven validation of the system of objectives to support developers of complex mechatronic systems [3]. The first phase serves to identify initial elements of the system of objectives that require data-driven validation. In the second phase, analysis use cases are defined by describing and prioritizing the technical system to ensure a uniform understanding of the analysis [22]. In the third phase the necessary data and analysis tools are identified. In the fourth phase, the required data and analysis tools are selected to lay the groundwork for the data analysis. The fifth phase focuses on evaluating the opportunities and risks of the selection. In the sixth phase, the actual data analysis is conducted. The results are evaluated, particularly assessing their significance, which necessitates close collaboration with the system's technology experts. In the final phase, decisions are derived from the analysis results. For a comprehensive decision, the analysis results must be shared with all relevant stakeholders for in-depth discussion. If the initial business question remains unanswered, previous phases can be repeated as necessary [3, 21].

3. Research Question and Methodology

Within the aim of enabling developers to conduct data analyses to validate the system of objectives, this work focuses on methodically support in the identification and selection of suitable data and analysis tools. To operationalize the aim of this work, the following research questions are formulated:

1. What are the influencing factors in the identification and selection of data and analysis tools in the data-driven development of complex mechatronic systems?
2. How can methodical support in the identification and selection of data and analysis tools in the data-driven development of complex mechatronic systems be developed?
3. What are the benefits of the methodical support in the identification and selection of data and analysis tools in the data-driven development of complex mechatronic systems?

This work is structured according to the Design Research Methodology (DRM) [23] and is carried out in the development department of a German machine tool manufacturer.

The descriptive study I (DS I) aims to answer the first research question by conducting nine interviews with data experts from the research environment for a better understanding of the influencing factors as well as the particularities of the research environment. The obtained knowledge is then used in the prescriptive study (PS) to develop methodical support for the identification and selection of data and analysis tools to answer the second research question. Finally, the developed solution will be evaluated in the descriptive study II (DS II). Therefore, workshops will be carried out to test the solutions with developers from the research environment in defined data analysis case studies.

4. Identification of Influencing Factors in the Identification and Selection of Data and Analysis Tools

Nine semi-structured guideline-based expert interviews are conducted to investigate the influencing factors in the identification and selection of data and analysis tools with participants from the Data Science and Data Engineering Domain. From the transcription of the interviews 182 core statements were summarized. The statements were then mapped to 15 factors that influence either the data or the tool identification and selection process, or both. (Fig. 1).

No. Of Personas Mentionings	Influencing Factor	Influence on	
		Data	Tool
9	Developer Expert Network	X	
6	Metadata Documentation	X	
8	Flexibility		X
7	Result Visualization		X
6	Performance		X
5	Frequency of Execution		X
4	Usability		X
4	Support		X
3	Result Provision		X
9	Developer Experience	X	X
9	Data Quality	X	X
8	Data Quantity	X	X
8	Data Pre-Processing	X	X
8	Access Rights	X	X
8	Interfaces	X	X

Figure 1. Influencing Factors on Data Identification and Selection

The **developer expert network** was derived as the first influencing factor. Experts are consulted by the participants regarding data identification, access, and selection. Thereby, three different parties can be identified: the data experts, data owners, and development domain experts. In the phase of data identification, data experts help to identify possible data sources. For more specific tasks like identifying specific parameters in a data source, the importance of including a development domain expert, meaning a person who developed the function or generated the data, is highlighted. For data access, the data owner, who is responsible for the data from a business perspective, must be contacted. For the data selection, all participants underline the importance of development domain knowledge.

Metadata documentation about data origin, access, meaning, and content was derived as a second influencing factor. ID5 highlights the independence from experts as a major advantage for an efficient data identification process. However, currently there is a lack of access to metadata. Furthermore, documentations suffer from a lack of standards which leads to large deviations in the level of detail as well as incomplete, duplicate, or outdated versions. According to ID1, these reasons can be traced back to a lack of central coordination and management of standards for metadata documentation.

4.1. Influencing Factors on Analysis Tool Identification and Selection

Flexibility was mentioned as a first influencing factor in analysis tool identification and selection. While some tools are more universal and can be used in a wide range of applications (e.g., programming languages) others are more specialized. Secondly, **result visualization** plays a central role. Depending on the complexity and purpose of the analysis, ID5 divides result visualization into 4 levels: a KPI, a diagram, a dashboard, or an advanced statistical model. **Performance** is another important factor to consider, especially for training complex statistical models and analyzing large data quantities. Some tools profit from using compute clusters for flexible scaling of computational power. Further, the **frequency of execution** of the respective analysis influences the process. The main differentiation is between analyses that are executed once or repeatedly. ID5 and ID6 highlight the capability of some tools to have automatic scheduling and update functions as well as maintenance capabilities while other experts point out future viability and long-term tool support. Lastly, the provided **result provision** is an important factor. ID2 mentions the importance to keep the target group of the analysis in mind since different stakeholders have different demands on the level of detail and the informative value.

4.2. Influencing Factors on Data and Analysis Tool Identification and Selection

Some of the factors influence both data and analysis tool identification and selection, like the **experience of the developer** conducting the analysis and working with specific data. Because of the complex nature of machine signals, exploratory data analysis (EDA) is often used to evaluate the data and refine the data selection. Further, all interviewed experts agreed on the influencing factor of **data quality** on the reliability of the analysis results. Before conducting the actual analysis, the experts attempt to evaluate the data in terms of completeness, accuracy, consistency, uniqueness, validity, and integrity with the help of EDA and make necessary adjustments. Aspects of interest thereby regard reasonable parameter limits, outliers, missing values, duplicates, and anomalies over time. The **data quantity** available and to be processed also plays a major role. ID6 and ID8 therefore both underline the importance of setting a data observation period for dynamic data as well as narrowing the data collection down to relevant attributes. **Data pre-processing** is required to bring the data into a format that suits the requirements of the tool. Other necessary pre-processing steps according to the experts include data aggregating, merging, or treating data quality. ID2 and ID5 herefor highlight the need for expertise in a programming language. The request for **access rights** and the long waiting time for approval (up to two weeks) are also recognized as an important factor. Without the necessary permissions, developers are currently not able to perform data identification techniques. A similar picture can be drawn for tool access with long waiting times, non-transparent permission regulations, and complex authorization processes.

As a final factor, **Interfaces** form the link between data and analysis tools. Depending on the platform, data sources and the complexity vary significantly. Three of the participants point out that often only certain data-tool combinations exist which are often not documented. Thus, the organization should focus on one integration platform as a central interface for easily linking various data from different areas with powerful, accessible, and easy-to-use analysis tools.

5. Development of the Methodical Support

With the developer expert network and the developer experience, two influencing factors were identified in the last chapter. Different developers have different expertise and experience and therefore require distinct levels of assistance in different stages of the process. Three distinct **areas of expertise** can be identified: the development domain expertise about the technical system, the experience with the relevant data, and the experience in conducting analysis. To develop a solution that provides optimal support to individual skillsets while still managing complexity and ensuring the applicability and efficiency of the model, five personas were developed. They cover the three categorys of experience with an individual level of expertise (basic, intermediate, expert). An overview of the personas categories is given in Fig. 2.

Area of expertise \ Persona	Nora New	Mike Monitor	Eric Experience	Dave Data	Alexa Analytics
Development domain	Intermediate	Intermediate	Expert	Basic	Intermediate
Data	Basic	Intermediate	Intermediate	Intermediate	Expert
Analytics	Basic	Intermediate	Basic	Expert	Expert

Figure 2. Overview of the personas and their level of expertise

Since transferring the personas to the organization creates a high degree of uncertainty, they were validated with the help of an online survey among 34 developers within the research environment. Based on its results and the received feedback, the personas were further adjusted and refined. It was also found that the first two categories of expertise are highly use case dependent. Managing this dynamic requires a self-assessment by the developer before each analysis use case. A metadata documentation was identified to be one of the central levers for securing easy data findability and understandability. Thus, a data catalog concept for metadata documentation is a central component of the solution and follows a data demand-driven approach. It hereby focuses on a three-level hierarchy model and fills identified knowledge gaps with the help of unconsidered metadata attributes. The **data object** serves as the main hierarchy level, for the description of data with metadata. A data object is defined as a description of a set of data, which correspond in their structure and meaning and is distinguishable by a unique identifier. A **data object group** hierarchically combines several data objects with similar characteristics in one group for increased overview in the data catalog. **Data object attributes** are characterized by the fact that they describe exactly one value. To ensure that the developed methodical support is easily accessible and usable by the developers it is important to link the single pieces together. Therefore, persona-dependent activity profiles were created guiding the developers through the process of data and tool identification and selection and providing the developed support as needed by the individual. Depending on the persona selected in each analysis use case, the developer has different mandatory and optional activities. For better clarity, the process was divided into three phases. The activities of these phases as well as the phase inputs and outputs will not be referred to in this work for scope reasons.

The first two sections of the template which are filled out in the previous phase previous phase, "Definition of Use Cases" serve as input for the first activity [22]. The first activity in phase 1 refers to the developer's **understanding of the metadata** concept. For this activity, a guide is provided to the developer. This activity has to be performed only once and thereafter can be skipped. As a second activity, the relevant **data needs to be identified** using the data catalog. To find the relevant data inside the catalog a systematic keyword search is conducted. After the data has been successfully identified, the developer needs to **pre-select the relevant data** for the use case. Therefore, metadata attributes like the "Description", "Available Products", "Example Data Value" and "Access Points" provide help. Now, the **analysis tool identification and selection** can be initiated. The analysis tool selection is dynamic and is performed for every analysis use case. Thus, the data catalog includes the metadata attribute „Required Analytic Skill Level", which suggests the minimum maturity level of the developer in the conduct of analyses based on the nature of the data. The three defined levels of this attribute correspond with the defined expertise levels in the conduct of analyses from the performed persona self-assessment. If the developer meets these minimum requirements, the use case specific requirements need to be considered. Some of the most important requirements include result visualization, and third-party result sharing but are not limited to them. If at this point, several different tools meet all requirements, the developer can choose one. The first phase is completed with a **gate check** where the third part of the use case template is filled out.

Phase 2 is about understanding how to request access to data and analysis tools. Data and tool access requests follow similar processes. To speed up the process, a rich set of metadata attributes help with a more precise pre-selection and standardized request processes to prevent unnecessary communication loops. The **gate** of the second phase is checked when access to data and tools is granted.

In the third phase, the data selection is reviewed with experts, and adjusted as needed. First, **a pre-analysis or EDA needs to be conducted**. A best practice is to document all assumptions and the questions that arise as a basis for discussion for the later expert meeting in the form of a presentation for better understanding. The mentioned aspects should then be **discussed in the expert meeting**. If additional data is identified that has not yet been included in the analysis, the data selection must be adjusted. This shows that the process of data and analysis tool identification and selection is iterative. In the last **gate check** the analysis use case template is updated with potential changes made to the data object selection and stored centrally for documentation purposes and as a reference for future analyses.

6. Benefit Evaluation of the Methodical Support

For the benefit evaluation, workshops were selected as a suitable method. 15 developers were given the same tasks selection in two similar analysis case studies in a realistic development scenario: once without the support and once with support. In the workshops it was of interest to investigate the differences between developers of different personas, how they vary in the problems they are facing, and their need for support. Therefore, three separate workshops were conducted, each investigating a group of five developers representing one persona of special interest. The first persona (Eric Experience) was picked because of the number of self-assignments by the developers in the conducted survey in the PS. To cover the persona, five Automation Developers took part. Nora New was picked since it has the highest deficits in experience and therefore is most dependent on support. This persona was represented by participants from the Modular Development and Product Management departments. Dave Data was selected because of his deficiency in the development domain and was represented by Data Scientists. The resulting data of the evaluation are shown in Fig. 3.

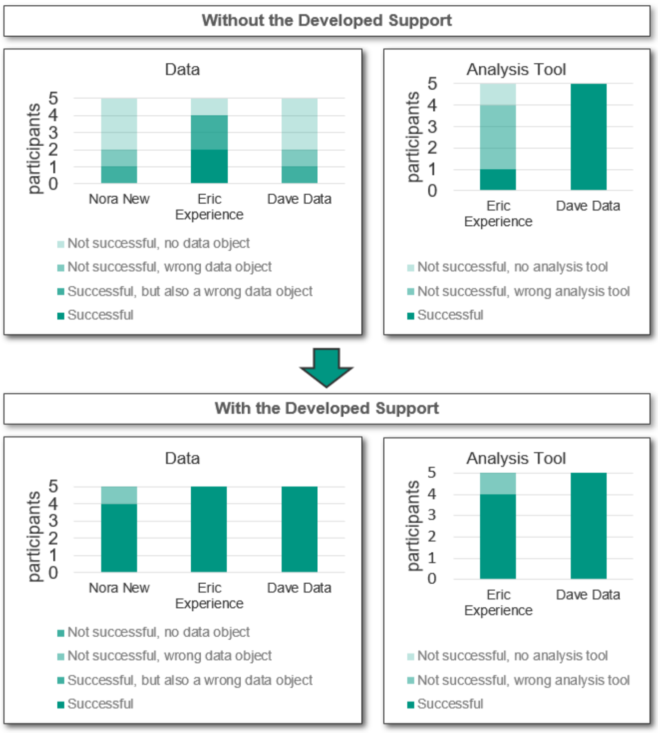

Figure 3. Evaluation results of the conducted workshops

In the Nora New workshop, there were problems meeting the time constraints. Therefore, in the second case study, the task regarding the selection of a suitable analysis tool was not answered by the participants.

In the first task of the first case study, developers were asked to identify and select the relevant data without the developed support of this work. Nora New, as well as Dave Data, showed a similar result distribution which implies that more data-affine data scientists also struggle to identify and choose the right data without support. Eric Experience already had more success in the first round. This is somewhat surprising since Eric Experience was not ranked higher in the experience with relevant data than Dave Data. This indicates that development domain expertise might also affect the experience with relevant data. Compared to this, with the developed support at hand, 14 of 15 developers were able to select the correct data. This shows a clear improvement for all personas. Factors of high relevance, for the developer's success, were identified to be the provision of example data records, information from which technical systems the data is generated, and the provision of descriptions on the data object attribute. All these features were elements of the developed data catalog and thus confirm its significance for successful data identification and selection.

The second task focused on the analysis tool identification and selection. Without the developed support at hand, six of the ten developers were able to select a suitable analysis tool. Significant differences between the different persona workshops could be seen. The Eric Experience developers had significant problems in choosing a suitable analysis tool which caused overwhelming and led to frustration among the developers. This was traced back to a lack of knowledge about available analysis tools existing in the organization and a lack of knowledge of important requirements and judgment of the tools' fulfillment of such requirements. Dave Data developers, on the other hand, had no problem choosing a suitable analysis tool even without the developed support at hand. This, however, was expected since they possess high experience in conducting analysis. With the developed support at hand four of five Eric Experience developers accomplished the task which shows a clear improvement. This was traced back to the provided step-by-step guidance of the developed support as well as its efficiency boost by directly providing skill requirements of the data at hand inside the data catalog. Like in the first round, Dave Data developers managed to select a suitable tool. However, their efficiency decreased from needing on average 34 percent of the given time without support to needing 68 percent because the received support material only cost them valuable time. In a real-life situation, this would have been prevented since the activity of analysis tool identification and selection provided in the activity profile of Dave Data is marked as optional for efficiency reasons. This again highlights the necessity of providing different levels of support in different phases to developers with different expertise profiles.

7. Discussion and Outlook

In this work, 15 factors were identified which influence the data or the analysis tool identification and selection. Different experts are required to conduct data analysis in complex mechatronic systems development which higlights the importance of interdisciplinary work. This is especially critical for beginners as they do not possess a network of experts yet. Three distinct areas of expertise were identified that must be covered for the successful execution of analysis use cases: the development domain expertise about the technical system, the experience with the relevant data, and the experience in conducting analysis. To develop a solution that provides optimal support while still managing complexity and ensuring the applicability and efficiency of the model, five personas were developed. Furthermore, metadata documentation was identified as crucial to decrease the complexity of technical data and the need of experts. Therefore, a data catalog concept was developed which focuses on the needs of end users to fill knowledge gaps with the help of specific metadata attributes. The developed support was finally linked together by developing persona-dependent activity profiles which guide the developers through the process. This ensures that developers get the necessary help they need while not being bothered by unnecessary, time-consuming support. The developed methodical support was finally evaluated by conducting workshops with 15 developer of different experience profiles. The developers were given the same tasks in two similar analysis case studies representing realistic development scenarios, once without and once with support. A significant improvement in success rate was achieved which confirms the added value of the developed methodical support.

The first limitation regards persona self-assignment. It was discovered that some people tend to overestimate or underestimate their competencies which can lead to persona misassignment. To ensure support by the model in such cases it was decided to still provide persona-irrelevant activities as optional activities. Further limitations concern the organizational environment. The concept is adapted to the present analytic maturity level of the organization. For example, the data infrastructure was relatively mature with a central data integration platform being available. Other organizations with other technology stacks, tools, interfaces, and processes might have to focus on other aspects and the support must be adapted to the specific environment.

Despite the proven benefits of the developed methodical support, there is still a need for further developments. The overall process model this work is part of must be continuously refined in iterative steps through further validation cycles. Furthermore, it is important to identify, observe, and support interfaces to adjacent processes inside the organization to ensure continuous improvement with the help of user feedback. Future work could also focus on refining data responsibility roles for technical and business users, developing maintenance strategies, and integrating metadata relevant to other parties besides end users, like data providers. Further, future work could focus on providing user feedback regarding data quality or developing a simple no-code analysis tool that can be used for data analysis with little to no expertise.

References

[1] F. Pfaff, S. Rapp, and A. Albers. Virtual-Reality-Umgebung für die Visualisierung von Entwicklungszielgrößen auf Basis des Referenzsystems im Modell der PGE – Produktgenerationsentwicklung in Entwerfen Entwickeln Erleben in Produktentwicklung und Design 2021, Dresden, Germany, 2021; pp. 175–187.

[2] F. Pfaff, S. Rapp, and A. Albers. Virtual-Reality-Umgebung für die Visualisierung von Entwicklungszielgrößen auf Basis des Referenzsystems im Modell der PGE – Produktgenerationsentwicklung in Entwerfen Entwickeln Erleben in Produktentwicklung und Design 2021, Dresden, Germany, 2021; pp. 175–187.

[3] E. Gholamzadeh Nabati and K. D. Thoben. On Applicability of Big Data Analytics in the Closed-Loop Product Lifecycle: Integration of CRISP-DM Standard. In: Product Lifecycle Management for Digital Transformation of Industries.: 13th IFIP WG 5.1 International Conference, PLM 2016, R. Harik, L. Rivest, A. Bernard, B. Eynard, and A. Bouras, Eds., Cham, Switzerland: Springer, 2016; pp. 457–467.

[4] S. Wagenmann, A. Krause, S. Rapp, A. Albers, L. Sommer, and N. Bursac. Application and Adaptation of a Process Model for the Data-Driven Validation of the System of Objectives. In 2022 IEEE International Symposium on Systems Engineering (ISSE), Vienna, Austria, 2022; pp. 1–8.

[5] S. Wagenmann, N. Bursac, S. Rapp, and A. Albers. Success Factors for the Validation of Requirements for New Product Generations – A Case Study on Using Field Gathered Data. In: Proceedings of the Design Society, vol. 2, pp. 1805–1814; 2022.

[6] A. Albers, Q. Lohmeyer, and B. Ebel. Dimensions of objectives in interdisciplinary product development projects. In: Proceedings of the 18th International Conference on Engineering Design (ICED11), vol. 2, pp. 256–265; 2011.

[7] A. Albers. Five Hypotheses about Engineering Processes and their Consequences. In: Proceedings of the TMCE, vol. 12, pp. 343–356; 2010.

[8] A. Albers, J. Reinemann, T. Hirschter, J. Fahl, and N. Heitger. Validation-Driven Design in the Early Phase of Product Development. Procedia CIRP, vol. 84, pp. 630–637; 2019.

[9] C. M. Eckert, T. Alink, and A. Albers. Issue driven analysis of an existing product at different levels of abstraction. In Proceedings of DESIGN 2010, Dubrovnik, Croatia, 2010; pp. 673–682.

[10] A. Albers, N. Bursac, and E. Wintergerst. Product generation development-importance and challenges from a design research perspective. In: Proceedings of INASE Conferences, pp. 16–21; 2015.

[11] A. Albers et al. Das Referenzsystem im Modell der PGE – Produktgenerationsentwicklung: Vorschlag einer generalisierten Beschreibung von Referenzprodukten und ihrer Wechselbeziehungen. KIT Scientific Working Papers, vol. 96; 2019.

[12] A. Albers, S. Rapp, C. Birk, and N. Bursac. Die Frühe Phase der PGE – Produktgenerationsentwicklung. In Stuttgarter Symposium für Produktentwicklung SSP 2017, H. Binz, B. Bertsche, and W. e. a. Bauer, Eds., Stuttgart, Germany: Fraunhofer Verlag, 2017; pp. 345–354.

[13] K. Miller. Data-Driven Decision Making: A Primer for Beginners. [Online]. Available: https://www.northeastern.edu/graduate/blog/data-driven-decision-making/ (accessed: Jan. 4 2023).

[14] C. Bange, M. Handford, and N. Janoschek. Information Culture: Leveraging the power of collective intelligence for better decision making. BARC Institute, Würzburg, Germany, 2014. Accessed: Dec. 12 2022. [Online]. Available: http://barc-research.com/research/information-culture/

[15] C. Anderson. Creating a data-driven organization. Beijing, China: O'Reilly; 2015.

[16] A. Labrinidis and H. V. Jagadish. Challenges and Opportunities with Big Data. In: Proceedings of the VLDB Endowment, vol. 5, no. 12, pp. 2032–2033; 2012.

[17] A. Gandomi and M. Haider. Beyond the hype: Big data concepts, methods, and analytics. In: International Journal of Information Management, vol. 35, no. 2, pp. 137–144; 2015.

[18] J. Hiba, H. Hadi, A. Hameed S., S. Hadishaheed, and A. Haji. BIG DATA AND FIVE V'S CHARACTERISTICS. In: International Journal of Advances in Electronics and Computer Science, vol. 2, no. 1, pp. 2393–2835; 2015.

[19] S. Lavalle, M. S. Hopkins, E. Lesser, R. Shockley, and N. Kruschwitz. Analytics: The New Path to Value: How the smartest organizations are embedding analytics to transform insights into action. [Online]. Available: https://sloanreview.mit.edu/projects/analytics-the-new-path-to-value/ (accessed: Jan. 5 2023).

[20] A. Díaz, K. Rowshankish, and T. Saleh, Why data culture matters. [Online]. Available: https://www.mckinsey.com/capabilities/quantum black/ our-insights/why-data-culture-matters (accessed: Jan. 5 2023).

[21] A. Albers, N. Burkhardt, M. Meboldt, and M. Saak. SPALTEN Problem Solving Methodology in the Product Development. In DS 35: Proceedings ICED 05, the 15th International Conference on Engineering Design, Melbourne, Australia, 2005; pp. 553–554.

[22] S. Wagenmann, F. Weidinger, M. Schöck, A. Albers and N. Bursac. Descriptive attributes of analysis use cases in the data-driven validation of elements in the system of objectives. In: Proceedings of the Design Society, pp. 1207-1216; 2023.

[23] C. Blessing. DRM, a design research methodology. London, United Kingdom: Springer; 2009.