Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 128 (2024) 13–18

www.elsevier.com/locate/procedia

34th CIRP Design Conference

Identification and Evaluation of Opportunities and Risks for Data-driven Validation of Systems of Objectives

Steffen Wagenmann[a]*, Rishab Arora[b], Artur Krause[c], Nikola Bursac[c], Albert Albers[a]

[a]*IPEK, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany*
[b]*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

Contemporary scientific research underscores the importance of data-driven validation within the product development process, particularly in relation to a system of objectives. Leveraging field-gathered machine data to inform decision-making in developing intricate mechatronic systems holds considerable promise. Despite the existence of various established models in the literature for data analysis, such as the widely adopted CRISP-DM process model for data mining, there remains a need for a specialized process model explicitly designed to assist developers involved directly in the product development process. This model should not only aid in conducting data analysis but also provide insights into the associated opportunities and risks inherent in the data analysis process. While it is tempting to focus solely on the potential benefits and opportunities offered by data analysis, it is equally essential to carefully consider the accompanying risks. Therefore, the objective of this study is to develop a systematic approach for evaluating both opportunities and risks in the data-driven validation of a system of objectives.

Keywords: Data Analytics; Data-driven validation; Design: challenges & Innovation

1. Introduction

The advancement of information technology and subsequent changes in the business and social landscape of digitalization have resulted in a significant increase in data volume across various industries. This growth consequently increased the demand and aspiration to derive value from available data assets [1]. However, recent development practices have demonstrated that developers face considerable challenges in actively utilizing customer data on machine usage for decision-making in mechatronic system development [2].

Data analytics provides a means to discover relevant information, structures, and patterns within data, facilitating the acquisition of new insights, identification of causality, prediction of future developments, and the generation of optimal decisions. Already available models include Knowledge Discovery in Databases (KDD) and the widely recognized Cross Industry Standard Process for Data Mining (CRISP-DM) support performing data analyses to derive valuable insights from data [3]. While the potential benefits of data analysis are compelling, it is imperative to acknowledge the associated risks e.g., deriving decisions based on false analysis results. Thus, conducting a detailed assessment of both opportunities and risks before initiating the data analysis process proves beneficial, as it offers a comprehensive understanding of the situation, particularly through evaluating all the opportunities and risks specific to a given analysis use case. However, in corporate practice, data analysis models often do not adequately incorporate these considerations, highlighting the need for a thorough investigation of this issue.

The challenge of assessing opportunities and risks in product development as well as the challenges inherent in effectively

10.1016/j.procir.2024.04.001

utilizing and leveraging data across various domains are well-acknowledged in scientific literature. Yet, practical applications that explicitly examine the opportunities and risks associated with data usage within the product development process of mechatronic systems are insufficient. Given the complexity of mechatronic systems, which integrate elements of various engineering disciplines into a cohesive unit, understanding the practical implications of data utilization necessitates direct investigations within the actual development processes of these systems. This approach is essential to capture the real-world intricacies and challenges faced by developers, thereby providing insights that are both relevant and applicable to the field to mitigate risks and capitalize on emerging opportunities.

2. Literature Review

2.1. The System Triplet in Product Development

The System Triplet of the Product Engineering model outlines product development as an ongoing interaction involving three interconnected systems: the *Operation System*, the *System of Objectives*, and the *System of Objects*, as illustrated in Figure 1. Iterative cycles drive the interaction between these systems, wherein a loosely defined system of objectives is progressively transformed into a well-defined system of objects through the operation system. Each cycle comprises two essential steps: creation and validation [4].

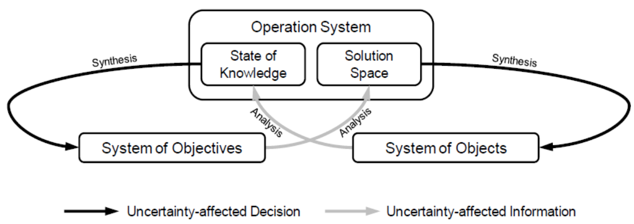

Figure 1. The System Triplet of Product Engineering

During the creation step, target analysis and object synthesis are performed, laying the foundation for the subsequent stages. The validation step encompasses object analysis and synthesis, which leads to the generation of knowledge. This knowledge is then employed to update the system of objectives, ultimately influencing the subsequent creation step. Consequently, this approach can be characterized as a validation-driven design, where the process is guided by the iterative validation and refinement of objectives [5].

2.2. Data-driven Validation in Product Development

The utilization of machine usage data for development decisions remains limited, despite the significant amount of data collected by numerous companies [6]. This inability to actively leverage such data for decision-making in development processes necessitates enabling developers to utilize usage data from technical systems for future decision-making [7]. The collected data contains valuable information, including patterns, dependencies, and trends, which can be utilized to enhance manufacturing processes and overall productivity. Consequently, Data Mining has emerged as a crucial tool for acquiring knowledge to improve decision-making [8]. As the volume of data grows exponentially, the ability to harness and interpret this data becomes both a challenge and an opportunity. The effective processing and structuring of vast data sets demand robust infrastructure, adequate resources, and specialized competence to extract meaningful insights and derive value.

Despite these challenges, the strategic utilization of data holds the potential to substantially improve decision-making capabilities. Especially in the early stages of product development, decisions are made under substantial uncertainty. The analysis of customer usage data serves as a validation instrument for system element objectives, offering an opportunity to enhance future system development and reduce market uncertainty. The systematic utilization of reference system elements for the data-driven validation of machine function usage holds substantial potential in the realm of product development [9].

2.3. Opportunity and Risk Analysis

Every project encounters uncertainties that can affect its success. It's important to recognize that some of these uncertainties may present opportunities while others could pose risks. Incorporating both risks and opportunities is vital in risk management because they can significantly impact project objectives. Neglecting either aspect can result in an incomplete understanding and suboptimal decision-making. Focusing solely on risks may lead to a defensive approach that overlooks potential benefits, while an exclusive focus on opportunities may lead to excessive risk-taking without adequate mitigation measures. Therefore, a comprehensive and balanced approach that considers both ensures a more scientifically rigorous and effective risk management process, optimizing overall performance and objective attainment. [10]. Hence, the data analysis process should not be exempt from evaluating opportunities and risks. Risk management is an iterative process involving four key activities: *identification*, *evaluation*, *mitigation*, and *control of risks*. These activities are iterative, as risks are typically not immediately and definitively eliminated. Thus, the impact of mitigation measures is monitored through risk control, followed by new risk identification due to the emergence of potential risks. All four activities are centered around a strategy component that orchestrates the risk management process [11].

Uncertainties can have a range of effects on project objectives, from complete failure to unexpected favorable outcomes. Nevertheless, conventional risk management processes tend to primarily concentrate on the potential adverse consequences of uncertainty. Consequently, significant efforts are dedicated to identifying and managing risks, while opportunities are often overlooked or addressed reactively, at best. Therefore, "uncertainty" serves as an overarching term that encompasses two categories: (1) "risks" solely referring to threats, i.e., uncertainties with negative effects, and (2) "opportunities," which denote uncertainties with positive effects. It is argued that opportunities and threats are not qualitatively distinct, as both involve uncertainty that can influence project objectives [10].

2.4. Process Model for data-driven validation of the system of objectives

Errors such as misinterpreting customer requirements during the definition of product specifications can lead to costly and time-consuming rework that possibly cannot be corrected with justifiable effort. Therefore, continuous validation is necessary especially in the early phase of product development, even if the degree of maturity is still low [12,13]. Given the significance of data-driven decision-making and the potential of modern big data applications, the manufacturing industry must consider these trends when developing future generations of complex mechatronic systems [14]. While existing literature offers various models for data analysis, with CRISP-DM being a prominent example, there remains a requirement for a process model specifically designed to provide direct assistance to developers engaged in product development activities. A process model fulfilling these requirements has been proposed by Wagenmann et al. [2]. In this proposed model, the key activities of CRISP-DM [15] are mapped to the respective elements of the SPALTEN problem-solving methodology [9]. The resulting iterative process model consists of seven distinct phases. The first phase involves identifying potential analysis objects representing requirements in the development process that must be validated. The second phase entails identifying technology experts with in-depth knowledge of the system and the respective data. The third phase involves identifying all required tools and data for the successful conduction of a data analysis. The fourth phase concerns the selection of all necessary tools and data for the analysis. The fifth phase involves identifying and evaluating potential opportunities and risks before conducting analytical activities, particularly assessing challenging analytical tasks with a high risk of failure but potentially significant value in their results. In the following sixth phase, all necessary procedures for data preparation and analysis are performed, where data can be temporarily validated based on the analysis results. The last phase establishes continuous knowledge management, ensuring systematic documentation that following analytical activities can benefit from the obtained knowledge during the former process [2,14].

3. Research Question and Methodology

Within the broader aim of enabling developers to conduct data analyses using field-gathered data and subsequently validate the system of objectives, this study focuses on enhancing developers' capabilities to evaluate opportunities and risks in the data-driven validation process. To operationalize the aim of this work, the following research questions are formulated:

(1) What are the opportunities and risks in a data analysis process for the data-driven validation of a system of objectives?

(2) How can the identified opportunities and risks be evaluated methodically for the data-driven validation of the system of objectives?

(3) What advantages and benefits are resulting from applying such a method for the validation of a system of objectives?

This work is structured based on the Design Research Methodology (DRM) by Blessing and Chakrabarti [16] and is illustrated in Figure 2.

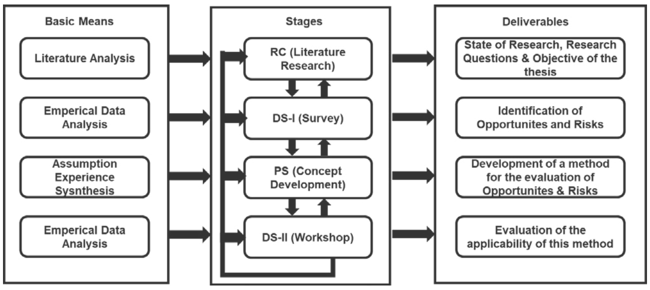

Figure 2. Design Research Methodology [17]

To answer the research questions, this study adopts a mixed-methods approach to explore opportunities and risks of conducting data analyses in the development process of complex mechatronic systems. This approach aims to ensure a robust and comprehensive assessment of the opportunities and risks encountered during the data analysis process. Initial efforts involve focused data analyses of specific use cases to identify current challenges and contextual complexities in the data analysis process for the validation of the system of objectives in the development of complex mechatronic systems. These analyses provide a singular perspective, recognizing that different analysts may encounter unique complexities in their respective contexts. However, these foundational analyses are crucial for setting the stage for a broader examination of the subject matter.

The subsequent phase includes a survey of 25 Data Analytics experts of the research environment with multiple years of experience in conducting data analyses based on field-gathered machine-usage data. This survey aims to gather insights on potential opportunities and risks associated with conducting such data analyses. To validate the survey results, two workshops are conducted, each with four participants, including developers and data analytics experts from the research environment. These workshops served as a critical review mechanism to validate the survey results, ensuring their reliability in the context of the product development process of complex mechatronic systems.

4. Results

4.1. Identification of Opportunities and Risks

The analysis of field-gathered machine-usage data can provide multiple opportunities for the data-driven validation of a system of objectives. The following Table 1 contains the opportunities identified in the research environment as well as the risks of conducting data analyses for the data-driven validation of a system of objectives identified in the research environment.

Table 1. Identified Opportunities and Risks through the Survey.

Survey No.	Opportunities	Survey No.	Risks
C1	High Customer Satisfaction	R1	Bad Decisions
C2	Improvement in Products	R2	Relevant data not found
C3	Optimization or Standardization of Processes	R3	No interface for data integration
C4	Competition advantage in the market	R4	Wrong analysis results
C5	Detection of vulnerabilities	R5	Not existing documentation
C6	Sound basis for decision making	R6	Intentional manipulation of data
C7	Better understanding of data	R7	Lack of technical understanding of the product
C8	Self-responsibility in development decisions	R8	Lack of data understanding
C9	New perspective for problems	R9	Lack of understanding of the data analysis process
C10	Agile working	R10	Requires Experts
C11	Regular growth of employees	R11	Unclear Analysis objectives
C12	Interdisciplinary Teamwork	R12	No motivation to perform data analysis
C13	Discovery of development potential	R13	Analysis too expensive/ time-consuming
C14	Cross-collaboration & dissolution of knowledge silos	R14	Analysis not informative enough
C15	Deep understanding of products	R15	Use of incorrect data to perform analysis
C16	Innovative Solutions	R16	Lack of training to perform complex analysis
C17	Improvement of tools infrastructure	R17	Difficult to estimate the time horizon at the beginning of an analysis

4.2. Evaluation of Opportunities and Risks

To evaluate the identified opportunities and risks during the data analysis process, a maturity model is developed and employed. A maturity model serves as a framework for assessing the level of maturity exhibited by entities such as organizations, individuals, or technologies. Maturity, in this context, refers to the ability to consistently improve in a specific discipline [17]. The maturity model illustrated in Figure 3, aims to depict real-world scenarios encountered by developers during data analyses in the research environment. The maturity model is characterized by a two-dimensional representation, wherein the X-axis represents the competence or maturity level of individual developers, and the Y-axis represents the maturity level of the organization. This two-dimensional portrayal provides a comprehensive view of the interplay between individual developer skills and the organizational context. To comprehensively represent the diverse scenarios encountered by developers during data analyses, it is imperative to consider both, the competence of the individual developer and the maturity level of the organization they are associated with.

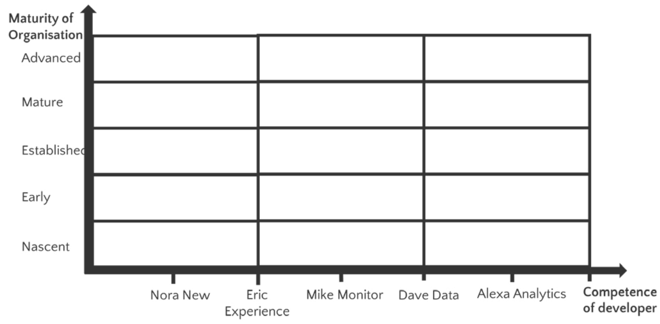

Figure 3. Maturity Model

By incorporating the competence of the developer and the organizational maturity level, the model aims to provide a holistic framework for assessing and understanding the complexities involved in data analysis processes. Based on competencies in the domain of development, data, and analytics, developers are classified into five personas: *Nora New*, *Eric Experience*, *Mike Monitor*, *Dave Data*, and *Alexa Analytics*. Each persona represents a maturity level from low (Nora New) to high (Alexa Analytics) based on the individual competence of a developer to conduct data analyses. Similarly, the maturity of the organization can be categorized into five distinct levels: *Nascent*, *Early*, *Established*, *Mature*, and *Advanced/ Visionary*. These categories are derived from differences observed across organizations in terms of five key dimensions: *Organization*, *Resources*, *Data Infrastructure*, *Analytics*, and *Governance* [17,18]. With the identification of opportunities and risks within the data analysis process and the established maturity model representing the resulting 15 different developer scenarios, an evaluation of opportunities and risks becomes feasible. With the conduction of two workshops, involving technical experts specializing in product development and data experts specializing in product analytics, the identified opportunities and risks are evaluated and mapped to the developed maturity model. To ensure an aggregated perspective, the following Figure 4 illustrates the mapped opportunities, wherein the position of each data point corresponds to the mean value derived from the combination of both conducted workshops.

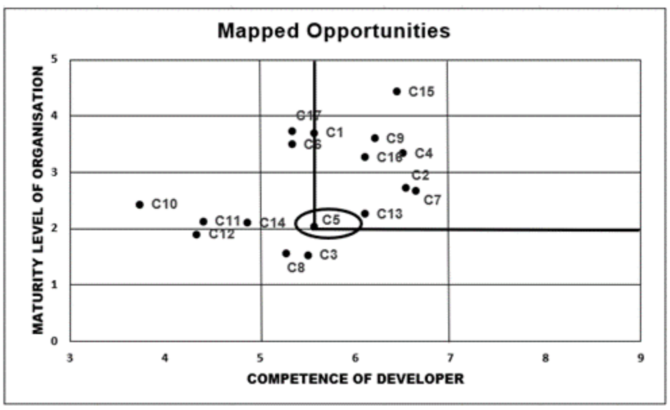

Figure 4. The mean of mapped Opportunities

To gain a comprehensive understanding of the mapped opportunities and their implications, opportunity C5: *Detection of Vulnerabilities*, is described in more detail. This opportunity is positioned between the personas of Eric Experience and Mike Monitor on the X-axis, aligned with the lower third maturity level. This positioning indicates developers represented by the personas situated to the right of this data point as Mike Monitor (6), Dave Data (7), and Alexa Analytics (8). Developers with similar competence levels have the potential to leverage this opportunity (C5) depending upon the maturity level of the organization rated at least at the third level *Established*. Notably, all individuals upper right should possess the capability to detect limitations after conducting the data analysis process. The mapping of risks follows a similar approach as the mapped opportunities, with one notable distinction. R10: *Requires Experts* pertain to the requirement of experts for conducting data analyses. In a mapped graph, this risk is positioned between the personas of Eric Experience (4) and Mike Monitor (6) on the X-axis, while being situated within the established maturity level of the organization on the Y-axis. This signifies that all developers possessing competencies below that of Mike Monitor (6), including Nora New (4), Eric Experience (5), and Mike Monitor (6), would encounter this risk during the data analysis process, based on that the organization's maturity level lies within the range of 1 (Nascent) to 3 (Established).

4.3. Dimensions of the Opportunities and Risks

By mapping each opportunity based on its dimensions, an integrated view of the opportunities across the entire maturity model can be created, showcasing the varying magnitudes and distributions within the data analysis process. In Figure 5 green regions indicate a higher likelihood of encountering numerous opportunities during the data analysis process, while red regions suggest fewer opportunities.

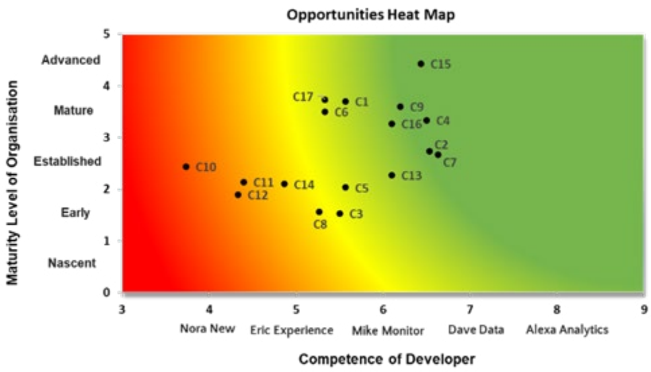

Figure 5. Magnitude of the mapped Opportunities Dimensions

This observation supports the notion that higher developer competence and organizational maturity enhance opportunities in data analysis. The colour differentiation offers insights into how opportunities are spatially distributed and how they relate to developer competence and organizational maturity. With agile working (C10), the regular growth of employees (C11), and interdisciplinary Teamwork (C12), a lower level of competence of the developer with a simultaneous lower maturity level of the organization provides fewer opportunities for organizational factors. Agile working practices promote collaboration, adaptability, and continuous improvement, which necessitate a certain level of competence and organizational maturity to fully leverage the benefits they offer. Regular growth of employees ensures a constant enhancement of skills and knowledge, enabling developers to address more complex tasks and seize opportunities effectively.

Additionally, interdisciplinary teamwork encourages the exchange of ideas and diverse perspectives, which are more likely to generate valuable opportunities for developers and organizations with higher competence and maturity levels. Therefore, developers and organizations must prioritize the development of competence, continuous growth, and interdisciplinary collaboration to maximize the opportunities available in their respective domains. Likewise, a higher level of competence of the developers and an increasing maturity level of the organization leads to several advantageous and amplifies opportunities, including improvement in Products (C2), a competitive advantage in the market (C4), a deep understanding of products (C15) as well as innovative solutions (C16). Improved competence allows organizations to identify and implement product enhancements, giving rise to higher quality and customer satisfaction. This enhances their competitive position in the market, providing distinct advantages. Additionally, increased competence enables organizations to develop a profound understanding of their products through their usage data, leading to the creation of innovative solutions that address customer needs and market demands. As the organization's maturity level grows, it gains access to more resources and supportive structures, facilitating product improvements, competitive positioning, and innovation. By continually enhancing competence and maturity, organizations can seize these opportunities and establish themselves as market leaders. By employing the same methodology, Figure 6 illustrates the representation of the identified risks and the magnitude of their dimensions.

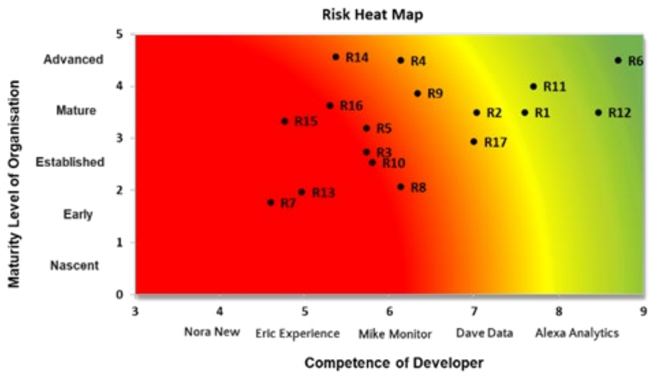

Figure 6. Magnitude of the mapped Risks Dimensions

In this visualization, each risk is connected to both the developer's competence on the x-axis and the organization's maturity level on the y-axis. The overlap of high-risk areas in this context indicates that these specific regions present significant challenges and the potential for negative outcomes in the data analysis process. Conversely, areas with minimal risk overlap, indicating lower levels of concern and fewer potential negative impacts. This differentiation emphasizes the importance of understanding and addressing risks based on their potential impact and the intersection of developer competence and organizational maturity. The lack of technical understanding about the product (R7) and the high cost or time-consuming data analysis (R13) pose significant challenges. Insufficient technical understanding may result in data misinterpretation and impair communication with stakeholders. The high cost or time requirements of analyses can result in delays, compromised depth, biased results, and missed opportunities for valuable insights. To address these challenges, organizations should prioritize gaining technical expertise, allocate adequate resources, and implement efficient analysis methodologies to ensure valuable outcomes.

5. Discussion and Outlook

The development of a systematic approach holds great potential for enhancing the effectiveness of data-driven validation in the development of complex mechatronic systems. By providing a structured framework, developers can navigate the complexities of data analysis more efficiently, leading to improved decision-making. However, the product development process entails unique challenges like integrating various data sources, addressing data quality problems, and maintaining data integrity during analyses. By defining the potential opportunities and risks naturally occurring during the product development process, developers are guided to evaluate risks more accurately even with limited relevant expertise in the data-driven domain. This requires both a correct assessment of the individual competencies of the developer as well as the maturity level based on the existing infrastructure, resources, and expertise of the organization regarding effective data utilization. While it is essential to highlight the potential benefits offered by data analysis, equal attention should be given to the potential risks and challenges. These may include biases in the analysis process and the potential for misinterpretation leading to incorrect conclusions. Understanding and mitigating these risks is crucial for ensuring the reliability and validity of the insights derived from the data. By correctly assessing the individual competencies as well as the maturity level of the organization regarding the utilization of data, further necessary steps can be identified, planned, and strategically executed to increase both. By categorizing these factors, individuals can better understand the risks involved and identify strategies for enhancing their competence. Elevating personal competence, in turn, contributes to organizational maturity, facilitating increased data availability and equipping the organization with the necessary tools and analytical capabilities. As competency and maturity levels rise, organizations can identify actions to better integrate and use advanced tools for data processing and analysis. Furthermore, strategically interpreting risks allows for the creation of countermeasures to effectively address potential challenges. This strategic assessment requires a careful balance between risk-taking willingness and the goal of exploiting opportunities, demanding a thoughtful approach to gain the benefits of data analysis while minimizing its risks. However, the scope of this study is limited by its focus on complex mechatronic systems within one specific research environment, potentially influencing the generalizability of its conclusions to other domains. Therefore, further research activities must be conducted in the domain of mechatronic system development but across different organizations considering varying degrees of complexity and more precise differences in the competency levels of developers to ensure broader applicability in evaluating opportunities and risks regarding data utilization in development practice. Additionally, different organizational structures, the maturity level, and data availability of other companies must be taken into account and examined. Despite these limitations, this research provides a comprehensive understanding of the opportunities and risks in data analysis within a specialized development context focusing on the development process of complex mechatronic systems.

References

[1] P. Haber, T. Lampoltshammer, M. Mayr. Data science - analytics and applications. In: International Data Science Conference (2. : 2019 : Salzburg): proceedings of the 2nd International Data Science Conference - iDSC2019. Wiesbaden: Springer Vieweg (Springer eBook Collection); 2019.

[2] S. Wagenmann, A. Krause, N. Bursac, S. Rapp, A. Albers, L. Sommer. Application and Adaptation of a Process Model for Data-Driven Validation of the System of Objectives; 2022.

[3] T. A. Runkler. Data Analytics – Models and Algorithms for intelligent Data Analytics. Third Edition; 2020.

[4] A. Albers, Q. Lohmeyer, B. Ebel. Dimensions of objectives in interdisciplinary product development projects. In: Proceedings of the 18th International Conference on Engineering Design; 2011.

[5] J. Reinemann, T. Hirschter, C. Mandel, J. Heimicke, A. Albers. Methodische Unterstützung zur Produktvalidierung in AR Umgebungen in der Frühen Phase der PGE – Produktgenerationsentwicklung, editors: D. Krause, K. Paetzold, S. Wartzack, editors. Design for X: Beiträge zum 28. DfX-Symposium; 2018.

[6] E. Bogner. Study Based Analysis on the Current Digitalization Degree in the Manufacturing Industry in Germany; 2016.

[7] S. Wagenmann, N. Bursac, S. Rapp, A. Albers. Success Factors for the Validation of Requirements for New Product Generations – A Case Study on Using Field Gathered Data. In: International Design Conference – Design; 2022.

[8] A. K. Choudhary, J. A. Harding, M. K. Tiwari. Data mining in manufacturing: a review based on the kind of knowledge; 2008.

[9] A. Albers. Verifikation und Validierung im Produkt-entstehungsprozess; 2016.

[10] D. Hilson. Extending the risk process to manage opportunities. In: Elsevier Science Ltd and IPMA; 2002.

[11] C. Ebert. Risikomanagement kompakt Risiken und Unsicherheiten bewerten und beherrschen. Springer-Verlag Berlin Heidelberg; 2013.

[12] A. Albers. Validation-Driven Design in the Early Phase of Product Development; 2019.

[13] A. Albers. PGE: Produktgenerationsentwicklung und Zukunftsvorausschau: Eine systematische Betrachtung zur Ermittlung der Zusammenhänge; 2018.

[14] S. Wagenmann, A. Krause, S. Rapp, S. Hünemeyer, A. Albers, N. Bursac. Process Model for the Data-driven Identification of Machine Function Usage for the Reduction of Machine Variants; 2022.

[15] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer, R. Wirth. CRISP-DM 1.0. Step-by-step data mining guide; 1999.

[16] Blessing, Lucienne T. M.; Chakrabarti, Amaresh. DRM, a design research methodology. Dordrecht, Heidelberg [u.a.]: Springer; 2009.

[17] F. Halper. TDWI Analytics Maturity Model. Assessment Guide; 2020.

[18] S. Wagenmann, J. Rall. Development of methods for systematic identification and selection of data and analysis tools in data-driven development of complex mechatronic systems; 2024.