



# Participative Method to Identify Data-Driven Design Use Cases

Simon Rädler and Eugen Rigger<sup>(✉)</sup>

V-Research GmbH, Dornbirn, Austria  
{simon.raedler,eugen.rigger}@v-research.at

**Abstract.** Paradigms such as smart factory and industry 4.0 enable the collection of data in enterprises. To enhance decision making in design, computational support that is driven by data seems to be beneficial. With this respect, an identification of data-driven use cases is needed. Still, the state of practice does not reflect the potential of data-driven design in engineering product development. With this respect, a method is proposed addressing the business and data understanding in industrial contexts and corresponding Product Lifecycle Management (PLM) environments. This allows to identify use cases for data-driven design taking into account business processes as well as the related data. In the proposed method, first the main process tasks are analyzed using a SIPOC analysis that is followed by a process decomposition to further detail and highlight corresponding applications using Enterprise Architecture principles. Following this, value stream mapping and design process failure mode effect analysis are used to identify sources of waste and the related causes. With this, a feature analysis of given data is proposed to identify use cases and enable to further use standard data science methods like CRISP-DM. The method is validated using the infrastructure of the Pilotfabrik at TU Vienna. The use case shows the applicability of the method to identify features that influence the cost of a product during the manufacturing without changing the functional specifications. The results highlight that different methods need to be combined to attain a comprehensive business and data understanding. Further, a comprehensive view of the processes is yielded that enables to further identify use cases for data-driven design. This work lays a foundation for future research with respect to data-driven design use cases identification in engineering product development.

**Keywords:** Data-driven-design · PLM · Enterprise architecture · ArchiMate · SysML

---

This work has been partially supported and funded by the Austrian Research Promotion Agency (FFG) via the “Austrian Competence Center for Digital Production” (CDP) under the contract number 854187.

© IFIP International Federation for Information Processing 2020  
Published by Springer Nature Switzerland AG 2020  
F. Nyffenegger et al. (Eds.): PLM 2020, IFIP AICT 594, pp. 680–694, 2020.  
[https://doi.org/10.1007/978-3-030-62807-9\\_54](https://doi.org/10.1007/978-3-030-62807-9_54)

## 1 Introduction

Approximately 70% of the product costs are determined during the design stage, making it crucial to come to informed decisions at this stage of product development [2,9,10]. In this respect, the use of computational methods and tools has shown to be an enabler to increase design performance [14] regarding efficiency and effectiveness of product development [29,33]. Nevertheless, state of practice still does not reflect the opportunities computational methods provide in the early stages of engineering design [28]. Still, recent trends such as Industry 4.0 [1,23,24] push enterprises towards implementation of smart factories that enable collection of data from Industrial Internet of Things (IIoT) [12] devices during the manufacturing process [7]. Furthermore, the advancement of mechanical engineering products towards smart cyber physical systems [37] due to the integration of IoT (Internet of Things) technologies provides information about the usage and status of a product in use [20]. The resulting sets of data from IIoT and IoT can be considered relevant data from the product lifecycle and can be used as an enabler for data-driven technologies in the design stage such as data-driven design. Data-driven refers to the application of computational methods that support decision-making in design based on data instead of intuition [5]. Hence, applications of data-driven design can rely on previous design revisions and related PLM data or other designs that feature similar characteristics. However, there is a lack of methods supporting the identification and formalization of use cases for data-driven design in the context of established product lifecycle management (PLM) strategies [32].

In response to this need, this work presents a method that enables the identification and formalization of data-driven design use cases in engineering companies based on participative workshops with the designers to analyze engineering processes and its supporting technological environments. It builds upon a PLM process decomposition by using Enterprise Architecture methods [4] tool and applies lean engineering methods for design process analysis [22]. In order to contextualize data artefacts with design features, a detailed analysis of the systems and data features is performed based on the Systems Modelling Language (SysML) [35]. The method is validated with a case study focusing on design and manufacturing of chess figures like in Fig. 1 using the infrastructure of the Pilotfabrik [15]. This paper is structured as follows: Sect. 2 analyzes the related literature with respect to existing case studies for data-driven design, methodologies to derive requirements for data mining and finally highlight research gaps in a summary. Section 3 proposes the new method which is evaluated in Sect. 4. Finally, a critical discussion with respect to validation and limitations is given which is finalized in Sect. 6 that highlights the conclusion and future work.

## 2 Background

In the following subsections, recent case studies with respect to data-driven design are reviewed. Following this, the state of the art for methodologies to

enable data-driven applications in engineering companies is analyzed. Next, an overview of modelling languages is given in order to support process and data mining. Finally, the state of the art is summarized and research gaps are highlighted.

## 2.1 Case Studies for Data-Driven Applications

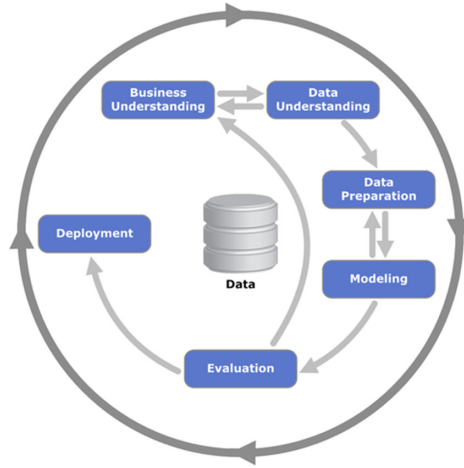
Recent scientific publications show different use cases according to the use of data-driven design with different level of details. The use cases comprise reports from variable selection over knowledge acquisition to decision modelling techniques. The data-driven design use case given in [18] shows a study that uses parameterized numerical simulation models and surrogate models to enable optimization based on a genetic algorithm. Since this approach not suggest a variable selection approach, the selection use case study of [17] is applicable. Here, all available high speed train design variables are analyzed each with respect to the relative importance based on previous engineering designers knowledge. For a more systematic approach to acquire the knowledge of a designer and formalize it for an impact estimation, [42] proposes to use focus groups studies. If the domain expert assignment of influence variables is not sufficient enough, an indication- and pattern-based method could be additionally used [13]. Since the data amount grows through the design process and data is rare at initial design steps, [21] recommends to start data-driven approaches in later design stages caused by the transformation from knowledge to data during the design process. The proposed framework support to numeralize different types of available variables and derive a decision tree based on subjective knowledge and previous data. As well, [8] suggests to build a decision tree based on geometrical behaviour to define the design space and reduce the complexity. The resulting decision tree builds a basis for rule based design tool support. A similar approach is given by [30], which focuses on minimizing the influence of uncontrollable (noise) variables. Here, a meta-heuristic is used to optimize the controllable variables to minimize the influence of the noise ones. The literature shows that case studies have been realized indicating the potential of data-driven design methods. However, these studies lack by using methodologies driving decisions with respect to data-driven design use cases, selection of data and traceability. Further, there is a lack of systematic evaluation of data-driven approaches in industrial practice based on metrics/key performance indicators.



**Fig. 1.** 3D Model of a bishop chess figure

## 2.2 State of the Art for a Data-Driven Methodology

An open standard process for data mining is Cross-industry standard process for data mining, known as CRISP-DM [40] is shown in Fig. 2. The methodology guides data mining project in industrial projects from the business understanding until the implementation of an application. More specifically tailored towards



**Fig. 2.** CRISP-DM standard process [27]

data science in the context of digital factories, the DMME builds an extension to CRISP-DM and proposes methods for data acquisition to detail the understanding of how machine, material, tools and product properties are related. Thereby, focus is put on identification and interrelation of different sources of data, e.g. sensor and machining data [16,39]. A similar approach for the usage of data from manufacturing can be found in [31]. The methodology proposes to use a failure mode analysis combined with a quality function deployment to pre-select influencing variables. However, the methodologies support on some points for data science projects while a comprehensive methodology for engineering design considering PLM, data back propagation and different types of sources is missing.

### 2.3 Modelling Languages

With respect to the process of CRISP-DM, business structure and processes needs to be collected to increase the business understanding. According to literature, enterprise architecture (EA) seems to be beneficial [19,34]. One of the main advantages of EA is the visual interrelation between the business and the application layer. However, the benefit of EA models is given, but a detail level selection is needed to reduce the amount of elements that needs to be drawn.

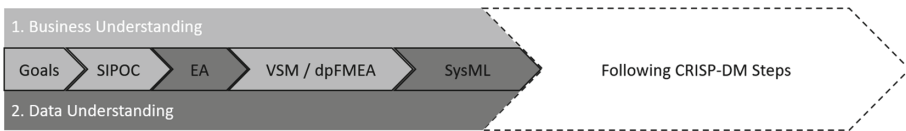
In order to model systems in the context of software applications, Systems Modeling Language (SysML) is widely used [35]. SysML is defined as an extension for Unified Modeling Language (UML) which is also an project by the object management group (OMG) [11]. One of the aims of SysML is to support the communication between interdisciplinary workers. Additionally, diagrams like the data-flow diagram supports in a validation process whether all needed interfaces are implemented.

### 2.4 Summary

The review of related literature shows that use cases are investigated for data-driven design applications in industrial contexts. However, the reviewed works do not highlight the applied methodology pursued for development of the use case. Hence, the presented use cases of data-driven design are difficult to trace and lack systematic evaluation of the potential in industrial contexts. In this respect, the application of methodologies to guide the integration and implementation of data-driven methods in design practice would lead to better comparability of published results and more comprehensive decision making in the evaluation of potential use cases for data-driven design. To reduce these shortcomings, this work further elaborates on the CRISP-DM methodology and refines it with respect to the details required for identification of use cases for data-driven design in established PLM environments. More specifically, focus is put on the formalization of the initial steps “business understanding” and “data understanding” of the CRISP-DM to provide the necessary context to comprehensively evaluate use cases for data-driven design. The refined steps apply methods stemming from enterprise architecture and systems engineering domains to leverage comprehensive analysis on both the (PLM) process as well as the system levels.

### 3 Method

In response to the needs highlighted in Sect. 2, this section proposes a participative method for identification of data-driven design use cases in engineering design while comprehensively taking the PLM into account with its technological environments as well as related data. The method builds upon the CRISP-DM methodology and extends its first two steps for business and data understanding as illustrated in Fig. 3 in order to make it applicable in an engineering design context.



**Fig. 3.** Method embedding in the CRISP-DM methodology

According to the figure shown in Fig. 3, the definition of goals is proposed as a first step to establish a business understanding and formalize the needs for design performance improvement. Further, goals are building the basis for the subsequent steps of identification and evaluation of data-driven design use cases. Next, one or multiple SIPOC analysis need to be conducted within participative workshops to define the scope of the investigated aspects of PLM. The workshops are conducted with engineers, designers and a workshop leader who

guides the re-engineering of the product lifecycle process steps and is familiar with the method that is presented in this paper. The yielded SIPOCS are then further refined using EA modelling to investigate the processes as well as the supporting technological environments and related data. Therefore, the relevant aspects of PLM can be comprehensively analyzed. In this respect, design process value stream mapping (VSM) [22] is applied to identify potential shortcomings that can be resolved using data-driven design, e.g. a lack of information back-flow. Additionally, design process failure mode effect analysis (dpFMEA) [6] is used to guide metrics derivation as proposed in [28]. Based on these sources of information, relevant data objects can be identified and contextualized with the initially identified goal. In particular, the SysML is used to link features of the data to the goals. Therefore, the necessary data understanding can be achieved which is required for the subsequent steps of the CRISP-DM methodology that address the mathematical modelling and identification of respective computational methods. In the following, the newly proposed steps are detailed.

### 3.1 Goals - Definition of Operative Goals

Once a system or a process to optimize is selected, a goal needs to be defined in order to guide the subsequent steps for identification of use cases for data-driven design. This is well aligned with existing approaches for metrics definition that state that goals need to be defined prior to selection of metrics and corresponding actions, e.g. the Goal-Question-Metric method [3]. Hence, goals can refer to specific design artefacts such as “improve lifetime of feature XY” or more generally to (parts of) the design process, e.g. “use less narrow tolerancing in detail design without losing functionality”. Additionally, the desired goal specifies whether data-driven design can rely on previous design revisions and related PLM data or other designs that feature similar characteristics.

### 3.2 SIPOC

The second step comprises the identification of the processes that are related to the previously identified goal. More particularly, all aspects of the product lifecycle that impact or are impacted by the investigated design artefact/process need to be comprehensively assessed. To acquire the knowledge about the related processes, a Supplier-Input-Process-Output-Customer (SIPOC) [41] is applied within participative workshops with the related engineers to gain a high-level overview of the processes and define the scope of investigations. Graphical modelling is applied to enable direct validation of the generated models by the participants. The SIPOC process captures the process in three to five main tasks (P), the related input (I) and output data (O). The main suppliers (S) and recipients (C) are connected considering read and/or write access with the input/output data. Figure 4 shows a generic example process with two tasks structured according to the SIPOC schema. Depending on the formalized goals, multiple SIPOC analysis might be needed to capture all related processes of the product lifecycle.

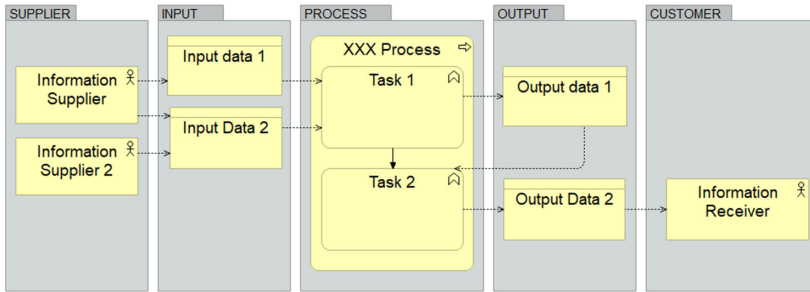


Fig. 4. SIPOC schema

### 3.3 EA - Detailed Process Modelling

The results from the SIPOC analysis are used as a basis to further detail the processes by decomposition of the tasks to yield single (design) activities [14]. In this respect, the ArchiMate modelling language [36] is applied to graphically model the specifics of the business processes, the related applications as well as the infrastructure including all the interrelations. The modelling expert guides the workshops by successive detailing of the main tasks from the SIPOC analysis based on a directed question-answer-talk. After modelling the design and business processes, the related applications are mapped to the activities. In particular, the main tasks are split up so to yield distinct design activities that are supported by a single application, e.g. “define initial 3D layout” that is supported by the CAD modelling application. Following this, the infrastructure needs to be modelled so that a comprehensive model of the enterprise architecture [19] is yielded. Particular focus is put on identification of data sources including its specific formats and accessibilities. Within this model, data artefacts and related tools can be directly contextualized with the business processes.

### 3.4 Value Stream Mapping/dpFMEA

To further strengthen the business understanding, information wastes are identified within the previously yielded model of the enterprise architecture using a design process value stream mapping [22]. Using a design process FMEA [6], effects and causes can be associated to the identified sources of information waste, e.g. redundant data generation that causes time to maintain and synchronize. This allows to derive related metrics [29] that can be used to validate the initially defined goal. In addition, the identified sources of waste permit to narrow down the amount of relevant data that needs to be investigated, e.g. when engineers identify unused information within artefacts.

### 3.5 SysML - Identification of Data Enabling Data-Driven Design

Based on the established business understanding and the yielded information about available data artefacts within the derived enterprise architecture model,

relevant sources of data can be identified so to potentially enable data-driven design. In particular, data attributes are identified and visualized using SysML block-definition diagrams [35]. Dependencies within the data can be highlighted based on data dependency interrelations that are used to focus explicitly on semantic dependencies within data objects, e.g. depth of a drilling hole and length of the selected drill. These data interrelations are visualized by adding information flows using the SysML “Item flow” relationship to the block-definition diagram. This data analysis based on decomposition and association of data artefacts to goals yields a comprehensive view of relevant data attributes. Hence, the data understanding is established. After an SysML expert modelled the diagram, another participative workshop is conducted to double check the results and to complete the interrelations. With the yielded results, influencing data sources can be identified to build the basis for systematic implementation of data-driven design. Building upon these findings, methodologies like CRISP-DM can be followed to support data preparation, modelling etc.

## 4 Validation

In this section, the method validation is presented with a use case taking design and manufacturing of bishop chess figures into account. The case study is conducted at the Pilotfabrik of the TU Vienna [15].

In the following, the results of each step are presented:

### 4.1 Goals - Definition of Operative Goals

The goal of the investigations for application of data-driven design is defined as to reduce the manufacturing time of a bishop chess figure during the turning process without changing its functional specifications or changing the chess figure shape. The desired data-driven design support has to affect the design attributes of a product and not the manufacturing parameters. Since the tolerance of a feature is not design specific, data is usable from other designs as well. Consequently, data is collected from multiple designs and revisions.

### 4.2 SIPOC

Since the Pilotfabrik corresponds to the prototype of a smart factory for demonstration scenarios, the supporting PLM processes such as the adaptive design process are not formalized. Hence, in this work, a generic design process according to [25] is assumed. Figure 5 indicates the corresponding SIPOC with In/Outputs and Suppliers and Customers as given in the context of the Pilotfabrik. The adaptive design process describes the optimization of an already existing design.



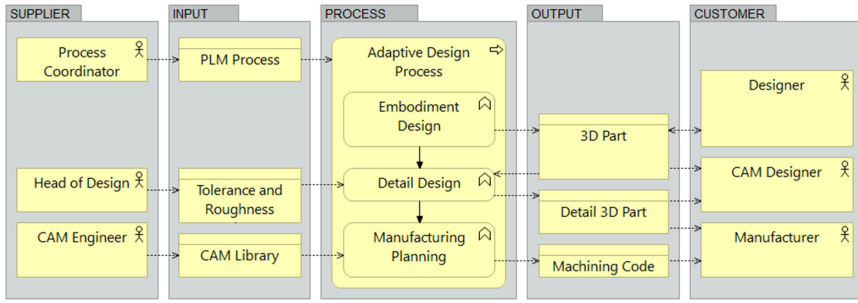


Fig. 5. SIPOC containing the main tasks of the adaptive design process

### 4.3 EA - Detailed Process Modelling

To accomplish a detail process model, first, a workshop was conducted visiting the shopfloor and analyzing the processes of a design expert. The main process tasks and the related tools were identified and the subjective worker view was modelled using a EA. Next, the process manager further detailed the EA model process tasks. The result of the two workshops yields in a EA model describing the current adaptive design process with the related application. Following this, the EA expert post processed the EA model with respect to readability and additionally annotated open questions, e.g. “Where is the CAD file stored?”. Finally, a workshop was conducted to answer the identified questions and to double check the model correctness. Figure 6 shows a subsection of the yielded workshops results. It shows the main design process tasks with its relation to the used applications and corresponding manufacturing data. The identified applications are SolidWorks for CAD, HyperMill for CAM and Centurio.works for data logging and orchestration using Business Process Model and Notation (BPMN) [26, 38] (blue) which are both stored on the network drive (green).

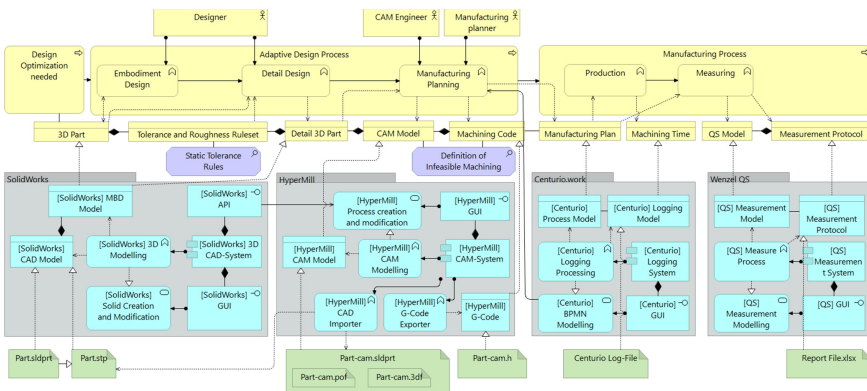


Fig. 6. Detail PLM process with respect to the SIPOCs (Color figure online)

#### 4.4 Value Stream Mapping/dpFMEA

The EA expert, the design expert and the process manager participated in a workshop to identify sources of waste using a value stream mapping. With respect to this, the EA expert used to ask questions and guided the workshop to yield the result that is visualized (purple) in the EA model Fig. 6. Thus, a dpFMEA was conducted during the workshop to identify causes and effects. The method to acquire the causes and effects was equal to the value stream mapping method, based on a question-answer-talk.

The result of the workshop yields in two sources of waste. The details according to the cause and effects are visualized in Table 1.

**Table 1.** dpFMEA result to identify further method steps

Activity	Failure mode	Effects	Causes
Detailing the design	Static tolerance rules	Increase in cost caused by narrow tolerances or roughness, time waste caused by missing tolerance selection support, additional iterations because of not manufacturable tolerances	Lack of tool support for dynamic tolerance and roughness rules
CAM programming	Definition of infeasible machining code	Inefficient machining code, additional iterations on failing machining code, no tolerance evaluation according to manufacturability	Lack of tool support for NC Code simulation

#### 4.5 SysML - Identification of Data Enabling Data-Driven Design

Based on the previous findings, the detail data understanding is improved during a analysis of data attributes. The findings of the dpFMEA is used to limit the amount of data objects from the EA process. The SysML model was created by analyzing the data object on a data format level by reviewing the API documentation of SolidWorks. The result was a comprehensive model on an attribute level without interrelations including possible redundancies. Next, a participative workshop was conducted to successively check attributes according to relevance and interrelations with other attributes. The result is presented in Fig. 7. The grey blocks are passive influencing the data-driven design. The yellow data block (dimension) is the object that mainly influences the desired data-driven design goal. The tolerance is influencing (red dashed arrow) the tool and corresponding turn feature specific parameters. The BPMN model and the QS model are both generated during the manufacturing task and are used for further implementation of the data-driven design support following further CRISP-DM steps.

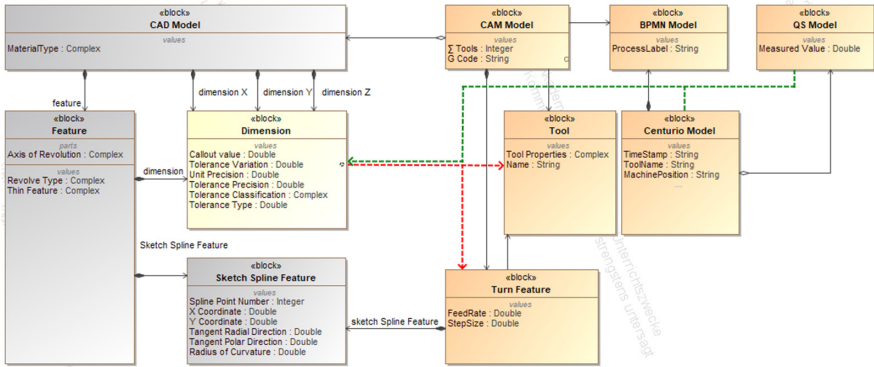


Fig. 7. System analysis to identify interrelations of feature (Color figure online)

## 5 Discussion

In this section the findings of this paper are discussed and assessed. The introduced method aims to be applicable in different engineering fields and validated based on a case study in the Pilotfabrik. The validation shows, that the method is valid for the identification and formalization of data-driven design use cases in mechanical engineering, more specific, with a turned part. The goal definition was beneficial to guide the participative workshops and sharpen the aim of the data-driven design approach. Additionally, a first idea of how the data-driven design use case might look like is given. The conducted SIPOC showed to be beneficial to share the business understanding between domain experts with different engineering background. During the workshops the graphical modelling of the main process tasks was helpful since all participants were able to follow the task modelling process and assess whether another process task is missing. Further, the EA modelling to create a more detailed view of the process and the related applications helped to introduce a first consolidated data understanding between the EA expert and the design expert. The question-answer-talk was useful to identify the detail tasks and related applications and the corresponding IT infrastructure. However, the modelling without EA templates caused difficulties in the level of detail which depends on the experience of the EA expert. The templates have to provide a generic process step with its corresponding application, functions and interface on the infrastructure level and a Information supplier and consumer on a business level. Next, the model proved to be a solid basis for the subsequent design process value stream mapping. The participants were able to identify information waste based on the diagram without further details that proved the aim of the EA model to increase the data understanding. Further, the dpFMEA was used to identify effects and causes with respect to the specific goal defined in method step one. This enabled to derive metrics to quantify the potential benefit of data-driven design and to narrow down the amount of relevant data that needs to be investigated in the following SysML step. With the

yielded result from the previous steps, the data understanding on a data level is improved during a data object decomposition with SysML. Feature of the data are identified with a data object analysis and, again, the visualization was beneficial to communicate between different domain experts. Still, a more detailed guidance of how to conduct the data object analysis with SysML is needed. Therefore, an abstraction of different data attributes needs to be introduced due to the possibly complicated interpretation of attributes which needs to be seen in context with other attributes. This work shows that a PLM processes decomposition is beneficial to generate a comprehensive view to identify data-driven design use cases. Corresponding, enterprises are able to introduce different applications that support the design process. From a scientific point of view, this work contributes by the introduction and evaluation of a new method to identify and formalize data-driven design use cases in engineering enterprises. More specific, knowledge is consolidated from different scientific communities like data science, engineering and lean management. The method contributes to industry by supporting the identification and formalization of data-driven design use cases while taking the experience of different domain experts into concern and giving the opportunity to derive metrics. Even though the validation highlights the beneficial effects of the method in a case study focusing on manufacturing of turned parts, the method requires further validation in different industrial contexts.

## 6 Conclusions

This paper contributes by presenting a new method to identify data-driven design use cases in engineering companies. The method is presented and validated with a case study in a smart factory with respect to design and manufacturing of turned chess figures. The method builds upon a systematic decomposition of related PLM processes using SIPOC analysis and enterprise architecture modelling to analyse business processes and the related infrastructure. To identify sources of waste and increase the business understanding, design process value stream mapping and dpFMEA are used. Additionally, a systematic analysis of systems and related data features is conducted using the graphical modelling language SysML. This supports the identification of data interrelations required to establish a profound data understanding to implement a data-driven method. The validation with experts with different engineering background that were involved in the case study shows that this method establishes a shared business and data understanding required to successfully identify and implement data-driven design in industry. Future work will focus on further industrial validation to validate the genericity of the approach for different domains and PLM strategies. Further, additional work needs to focus on development of templates and supporting tools to enhance usability of the method.

## References

1. Andelfinger, V.P., Hänisch, T. (eds.): *Industrie 4.0: Wie cyber-physische Systeme die Arbeitswelt verändern*. Gabler Verlag (2017). <https://doi.org/10.1007/978-3-658-15557-5>
2. Barton, J.A., Love, D.M., Taylor, G.D.: Design determines 70% of cost? A review of implications for design evaluation. *J. Eng. Des.* **12**(1), 47–58 (2001). <https://doi.org/10.1080/09544820010031553>
3. Basili, V.R., Caldiera, G., Rombach, H.D.: *The Goal Question Metric Approach* (1994)
4. Bernus, P., Laszlo, N., Schmidt, G. (eds.): *Handbook on Enterprise Architecture*. Springer, Heidelberg (2003). <https://doi.org/10.1007/978-3-540-24744-9>
5. Brynjolfsson, E., Hitt, L.M., Kim, H.H.: Strength in numbers: how does data-driven decision making affect firm performance? *SSRN Electron. J.* (2011). <https://doi.org/10.2139/ssrn.1819486>
6. Chao, L.P., Ishii, K.: Design process error proofing: failure modes and effects analysis of the design process. *J. Mech. Des.* **129**(5), 491 (2007). <https://doi.org/10.1115/1.2712216>
7. Chen, B., Wan, J., Shu, L., Li, P., Mukherjee, M., Yin, B.: Smart factory of Industry 4.0: key technologies, application case, and challenges. *IEEE Access* **6**, 6505–6519 (2018). <https://doi.org/10.1109/ACCESS.2017.2783682>
8. Du, X., Zhu, F.: A new data-driven design methodology for mechanical systems with high dimensional design variables. *Adv. Eng. Softw.* **117**, 18–28 (2018). <https://doi.org/10.1016/j.advengsoft.2017.12.006>
9. Ehrlenspiel, K., Kiewert, A., Lindemann, U.: *Cost-Efficient Design*. Springer, Heidelberg (2007). <https://doi.org/10.1007/978-3-540-34648-7>
10. Ferreirinha, P., Hubka, V., Eder, W.: Early cost calculation - reliable calculation, not just estimation. In: *Design for Manufacturability*, ASME DE, vol. 52, pp. 97–104 (1993), <https://www.researchgate.net/publication/279998839>
11. Fisher, D.: *OMG | Object Management Group* (1989). <https://www.omg.org/Library/Catalog:www.omg.org>
12. Gilchrist, A.: Introduction to the industrial Internet. In: *Industry 4.0*, pp. 1–12. Apress, Berkeley (2016). [https://doi.org/10.1007/978-1-4842-2047-4\\_1](https://doi.org/10.1007/978-1-4842-2047-4_1)
13. Gröger, C., Niedermann, F., Mitschang, B.: Data mining-driven manufacturing process optimization. In: *Proceedings of the World Congress on Engineering 2012 Vol III*, vol. 3 (2012)
14. Haffey, M., Duffy, A.: *Design process performance management support* (2003)
15. Hennig, M., Reisinger, G., Trautner, T., Hold, P., Gerhard, D., Mazak, A.: TU Wien Pilot Factory Industry 4.0. *Procedia Manuf.* **31**, 200–205 (2019). <https://doi.org/10.1016/j.promfg.2019.03.032>
16. Huber, S., Wiemer, H., Schneider, D., Ihlenfeldt, S.: DMME: data mining methodology for engineering applications - a holistic extension to the CRISP-DM model. *Procedia CIRP* **79**, 403–408 (2019). <https://doi.org/10.1016/j.procir.2019.02.106>

17. Jiang, J., Ding, G., Zhang, J., Zou, Y., Qin, S.: A systematic optimization design method for complex mechatronic products design and development. *Math. Probl. Eng.* **2018**, 1–14 (2018). <https://doi.org/10.1155/2018/3159637>
18. Knight, D.: Data driven design optimization methodology a dynamic data driven application system. In: Sloot, P.M.A., Abramson, D., Bogdanov, A.V., Gorbachev, Y.E., Dongarra, J.J., Zomaya, A.Y. (eds.) ICCS 2003. LNCS, vol. 2660, pp. 329–336. Springer, Heidelberg (2003). [https://doi.org/10.1007/3-540-44864-0\\_34](https://doi.org/10.1007/3-540-44864-0_34)
19. Lankhorst, M., Proper, H., Jonkers, H.: The anatomy of the archimate language. *Int. J. Inf. Syst. Model. Des.* **1**(1), 1–32 (2010). <https://doi.org/10.4018/jismd.2010092301>
20. Lee, J., Bagheri, B., Kao, H.A.: A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manuf. Lett.* **3**, 18–23 (2015). <https://doi.org/10.1016/j.mfglet.2014.12.001>
21. Liu, C., Chen, X.: Data-driven design paradigm in engineering problems. *Proc. Inst. Mech. Eng. Part G J. Aerosp. Eng.* **231**(8), 1522–1534 (2017). <https://doi.org/10.1177/0954410016653502>
22. Mcmanus, H.: Product Development Value Stream Mapping (PDVSM) Manual Release 1.0 (2005). [https://www.semanticscholar.org/paper/Product-Development-Value-Stream-Mapping-\(PDVSM\)-Mcmanus/5349ab84748e957cd05c9dfe5eede8307c07be02](https://www.semanticscholar.org/paper/Product-Development-Value-Stream-Mapping-(PDVSM)-Mcmanus/5349ab84748e957cd05c9dfe5eede8307c07be02)
23. Müller, J.M., Voigt, K.I.: Sustainable industrial value creation in SMEs: a comparison between Industry 4.0 and made in China 2025. *Int. J. Precis. Eng. Manuf. Green Technol.* **5**(5), 659–670 (2018). <https://doi.org/10.1007/s40684-018-0056-z>
24. Mutlu, B., Yamaoka, F., Kanda, T., Ishiguro, H., Hagita, N.: Nonverbal leakage in robots: communication of intentions through seemingly unintentional behavior. In: Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction - HRI 2009, La Jolla, California, USA, p. 69. ACM Press (2009). <https://doi.org/10.1145/1514095.1514110>
25. Pahl, G., Beitz, W.: Pahl/Beitz Konstruktionslehre Methoden und Anwendung erfolgreicher Produktentwicklung. Springer, Heidelberg (2013). <https://doi.org/10.1007/978-3-642-29569-0>
26. Pauker, F., Mangler, J.: centurio.work - higher productivity through intelligent connectivity. In: Wiener Produktionstechnik Kongress, vol. 4. New Academic Press OG (2018)
27. Riepl, W.: CRISP-DM: Ein Standard-Prozess-Modell für Data Mining (2012). <https://statistik-dresden.de/archives/1128>. Library Catalog: statistik-dresden.de
28. Rigger, E., Vosgien, T.: Design automation state of practice - potential and opportunities. In: DS 92: Proceedings of the DESIGN 2018 15th International Design Conference, pp. 441–452 (2018). <https://doi.org/10.21278/idc.2018.0537>. <https://www.designsociety.org/publication/40462>
29. Rigger, E., Vosgien, T., Shea, K., Stankovic, T.: A top-down method for the derivation of metrics for the assessment of design automation potential. *J. Eng. Des.* 1–31 (2019). <https://doi.org/10.1080/09544828.2019.1670786>
30. Sadati, N., Chinnam, R.B., Nezhad, M.Z.: Observational data-driven modeling and optimization of manufacturing processes. *Expert Syst. Appl.* **93**, 456–464 (2018). <https://doi.org/10.1016/j.eswa.2017.10.028>

31. Stanula, P., Ziegenbein, A., Metternich, J.: Machine learning algorithms in production: a guideline for efficient data source selection. *Procedia CIRP* **78**, 261–266 (2018). <https://doi.org/10.1016/j.procir.2018.08.177>
32. Stark, J.: PLM implementation strategy and plan. *Product Lifecycle Management (Volume 2)*. DE, pp. 555–565. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-24436-5\\_29](https://doi.org/10.1007/978-3-319-24436-5_29)
33. Stjepandić, J., Wognum, N., Verhagen, W.J.C. (eds.): *Concurrent Engineering in the 21st Century: Foundations, Developments and Challenges*. Springer, Cham (2015). <https://doi.org/10.1007/978-3-319-13776-6>
34. Tamm, T., Seddon, P.B., Shanks, G., Reynolds, P.: How does enterprise architecture add value to organisations? *Commun. Assoc. Inf. Syst.* **28** (2011). <https://doi.org/10.17705/1CAIS.02810>
35. The Object Management Group: *OMG SysML Home | OMG Systems Modeling Language* (1989). <https://omgsysml.org/>
36. The Open Group: *ArchiMate® 3.1 Specification* (2019). <https://pubs.opengroup.org/architecture/archimate3-doc/>
37. Wang, L., Törngren, M., Onori, M.: Current status and advancement of cyber-physical systems in manufacturing. *J. Manuf. Syst.* **37**, 517–527 (2015). <https://doi.org/10.1016/j.jmsy.2015.04.008>
38. White, S.A., Miers, D.: *BPMN modeling and reference guide: understanding and using BPMN; develop rigorous yet understandable graphical representations of business processes*. Future Strategies Inc., Lighthouse Point, Fla (2008)
39. Wiemer, H., Drowatzky, L., Ihlenfeldt, S.: Data mining methodology for engineering applications (DMME) - a holistic extension to the CRISP-DM model. *Appl. Sci.* **9**(12), 2407 (2019). <https://doi.org/10.3390/app9122407>
40. Wirth, R., Hipp, J.: CRISP-DM: towards a standard process model for data mining. In: *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining* (2000)
41. Yang, K., El-Haik, B.S.: *Design for Six Sigma: A Roadmap For Product Development*, 2nd edn. McGraw-Hill, New York (2009)
42. Zhang, J., Ding, G.F., Zhou, Y.S., Jiang, J., Ying, X., Qin, S.F.: Identification of key design parameters of high-speed train for optimal design. *Int. J. Adv. Manuf. Technol.* **73**(1–4), 251–265 (2014). <https://doi.org/10.1007/s00170-014-5822-7>