Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 109 (2022) 580–585

www.elsevier.com/locate/procedia

32nd CIRP Design Conference

Business-to-Analytics Canvas – Translation of Product Planning-Related Business Use Cases into Concrete Data Analytics Tasks

Melina Panzner[a]*, Maurice Meyer[a], Dr. Sebastian von Enzberg[b]
Prof. Dr.-Ing. Roman Dumitrescu[a,b]

[a]*Heinz Nixdorf Institute, University of Paderborn, Fürstenallee 11, 33102 Paderborn, Germany*
[b]*Fraunhofer Institute for Mechatronic Systems Design, Zukunftsmeile 1, 33102 Paderborn, Germany*

* Corresponding author. Tel: +49-5251-5465-346. *E-mail address:* melina.panzner@hni.uni-paderborn.de

Abstract

Cyber-physical systems (CPS) generate huge amounts of data during the usage phase. By analysing these data, CPS providers can systematically uncover hidden product improvement potentials for future product generations. The successful implementation of such analytics use cases depends to a large extent on whether the stakeholders involved succeed in coordinating their goals and procedures. In particular, product managers and data scientists must come to a common understanding in the context of defining and concretizing the use cases. A common vocabulary is necessary so that the data scientists or those responsible for analysis can determine target-oriented, analysis-capable use cases with which the processing of the data can start quickly and successfully. The research question that arises at this point is: How can business goals or use cases be translated into realizable analytics use cases or tasks? In this paper we present the Busines-to-Analytics Canvas as a result of an action design research approach. It supports the translation of business use cases and goals into concrete data analytics tasks for product planning. By providing various information elements and guiding questions, the canvas helps data scientists translate the business goal into a data analytics approach, i.e., an algorithm class, and gather the necessary information to start processing data.

© 2022 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 32nd CIRP Design Conference

Keywords: Data Analytics, Use Cases, Translation, Business Goals, Data Analytics Tasks, Product Planning

1. Introduction

Recent technical developments enable the systematic collection and analysis of huge amounts of field or usage data from cyber-physical systems (CPS). The insights generated from data analytics can reveal potentials for product improvements and help manufacturers of those CPS to optimize product performance and better adapt to actual customer needs [1]. This is an important task in the early phases of product development such as product planning [2]. As a general framework, standard models like CRISP-DM show the relevant process steps of data analytics projects: business understanding, data understanding, data preparation, modeling, evaluation, and deployment [3]. According to CHAPMAN and CLINTON, the first phase, business understanding, involves transforming the business goal (e.g. "Increase catalog sales to existing customers") into a data analytics goal ("Predict how many widgets a customer will buy, given their purchases over the past three years, demographic information and the price of the item") [4]. A data analytics goal comprises the outputs that enable the achievement of the business objectives. For this transformation business managers and data scientists need to work and communicate closely to ensure that business goals are clearly defined and fit with analytics activities. However,

2212-8271 © 2022 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 32nd CIRP Design Conference
10.1016/j.procir.2022.05.298

This is a resupply of March 2023 as the template used in the publication of the original article contained errors. The content of the article has remained unaffected.

problems often arise at this intersection [5]. When using data analytics in product planning, the use case definition and derivation of data analytics tasks is also an important and challenging step [6].

While product managers often have a clear understanding of their strategic goals (e.g., identifying the weak points and causes to improve the product), they are not clear about how data-driven analytics methods can help achieve those goals and formulate appropriate use cases and questions or hypotheses. On the other hand, data scientists are often stuck in their local view, deep in the algorithms, and struggle to ask the questions necessary to concretize the use case in more detail. In summary, they speak different languages with regard to their goals and procedures. The research questions this raises for us are: How can product planning-related business goals or use cases be translated into realizable analytics use cases or tasks? What needs to be concretized, i.e., which information and parameters are necessary to be able to start quickly with data processing?

The goal of this paper is to develop a visual collaboration tool that helps product managers and developers as well as data analysts to gather the necessary information to deepen the business use case, which can be used to derive one or more possible analytics approaches in the end. Especially the determination of a rough solution approach is to be regarded as an important milestone, since it gives orientation for the selection and specification of the following tasks.

The paper is structured as follows: The subsequent section discusses related work about methods and tools for defining business and data analytics use cases. In section 3, the used research methodology is explained. The research process including results is presented in section 4. This is followed by an application example in section 5, before the paper ends with a conclusion and an outlook.

2. Related Work

There are few approaches to support the task of defining a data analytics use case in connection with the business perspective.

As already mentioned, the CRISP-DM standard helps to structure data mining projects [3]. However, it does not suggest any methodological process for this purpose.

CHAPMAN ET AL. propose some tasks, outputs and activities in the step of determining data mining goals in their step-by-step data mining guide [4]. However, they do not get very specific about how the process of translation can succeed, for example.

NALCHIGAR and YU propose a modeling framework for requirements analysis and design of data analytics systems [5]. It consists of a business view, an analytics design view and a data preparation view. These views are linked together to connect business strategies to analytics algorithms and data preparation activities. The conceptual modeling framework provides holistic modeling support that connects business goals to advanced analytics system design but does not support the process of translating the business goals into technical terms concretely and in an easy methodological way.

MARBAN ET AL. propose a methodological approach to guide the development of a business model of the decision-making process within an organization [7]. The business model is used to get organizational requirements and use cases applied to data mining projects. The business decision-making model is translated into use cases based on heuristics. However, developing the models requires time and experience with modeling techniques.

MARBAN and SEGOVIA present an extension of the UML modeling language for data mining projects (DM-UML) covering the documentation needs for a project conforming to CRISP-DM [8]. This can be used as tool for modeling and connecting the business understanding with modeling phase and other phases of an analytics project. They differentiate a data mining use case model, where data mining use cases are obtained from the business use cases and goals (business analysis model). Together with the use cases, the data mining objectives for each use cases must be established in the data mining goal model. The goals are established in terms of the business goals as they should be and are a translation of the business problem to problems expressed in data mining terms, such as cluster data, create a predictive model. The necessary relationship between data mining goals and business goals is considered in the model but does not support the implementation of the translation.

Another method is the data analytics canvas, a semi-formal specification technique for describing an analytics use case and the necessary data infrastructure during the early planning and specification of an analytics project [9]. It can be used for communication of all stakeholders involved in the realization of an analytics project. The Analytics Canvas provides the framework for a data science project, but does not supply tool support for the task of translating the business goal into a data analytics task.

Consequently, there is a need for a simple method that supports the data scientist in the translation and technical concretization of the business use case.

3. Methodology

This research applies an Action Design Research (ADR) project to investigate how data scientists proceed when they are confronted with a business goal, in order to collect the important information for the concretization of a use case. Design science research or ADR are an established scientific framework for design-oriented research in the context of information systems [10, 11]. ADR gives the opportunity to empirically study the present research topic in a real business context. It also fosters collaboration and interaction with practitioners [12]. Table 1 gives an overview of the different stages of the research project, following a structure proposed by Sein et al. [10]. These are 'problem formulation', 'building, intervention and evaluation', 'reflection and learning' and 'formalization of learning' In the problem formulation stage, the to-be-solved problem is described based on a literature review and a first research from us on this topic. Next in the second phase the problem framing is used to design an artifact in an iterative way. Here the workshop concept is built and conducted with practitioners. The third stage, reflection and learning, ensures, on the one hand, that knowledge contributions are identified and, on the other hand, that the research process is adjusted based on evaluated preliminary findings.

This is a resupply of March 2023 as the template used in the publication of the original article contained errors. The content of the article has remained unaffected.

Participants were observed during the workshop and furthermore interviewed afterwards. Finally, the fourth stage of the ADR process is focused on the formalization of learning. This stage is characterized by inductive reasoning based on the individual findings and observations made during the prior stages. Hence, formalization of learning focusses on the generalization and abstraction of the results from the implementation of the translation process on a few specific use cases, which includes the development of general learning principles. Finally, these principles were consolidated and structured into a canvas for general transformation of product planning business use cases into data analytics use cases. Applicability is demonstrated in a follow-up case study. The implementation and results of the research process are described in the next section.

Stage	**Outcome**	**Methods**
Problem formulation	problem description	literature review, workshop study
Intervention	performed “translation process” with practitioners	workshop
Reflection and learning	knowledge collection	observation, discussion
Formalization of learning	generalization of the results into a canvas	

Table 1: Stages of ADR research approach

4. Research process and results

4.1. Problem Formulation

As already mentioned, the translation of the business use case into data analytics use cases/tasks is part of the standard procedure for data mining projects. The fact that methodological support is necessary for this is shown, among other things, by the authors in section 2. We attempted to explore what this might look like in the context of product planning in a previous research. Here we tried to answer the research question “How can data analytics use cases in product planning be specified and translated into concrete analytics tasks?” [13]. The results of the study showed a clear need for supporting methods and tools for defining and specifying analytics use cases in product planning. In the following we summarize key challenges and requirements and potential solutions.

- for data scientists, data analytics questions formulated by the product experts (e.g., ‘Which component triggers an error?’) do not provide enough information to derive a solution approach. The product experts need to provide further details like goals and problems, e.g., by using a questionnaire or a checklist. A more standardized question format could also be investigated to facilitate the assignment of an algorithm class. Templates for formulating questions could be helpful in this regard.
- The creation of suitable, analysis-enabling sub-questions seems to be a critical factor, but not very intuitive for data scientists.
- Rather, data scientists tend to focus on algorithm classes from the beginning. They translated the questions into analytics problems using algorithm classes such as clustering or classification.

4.2 Building, Intervention & Learning

Based on these findings, we have developed a new workshop concept to get new learnings how data scientists proceed when they receive a concrete "job" with a question to be answered, then defining the data analytics goal and tasks and what are critical aspects to consider. According to this concept, goals and related questions, defined by product managers or business managers, are to be used as input. These now need to be specified. For this purpose, the questions should first be assigned to an algorithm class. The possible algorithm classes are predefined. In parallel, assumptions, questions and reformulations are to be recorded that are necessary for the assignment. In the next step, further aspects are to be collected that serve to deepen the use case and enable the start of processing.

The workshop was conducted in two iterations. Participants of the workshop were 7 scientific research associates in the area of Industrial Data Science and the associated department head (senior expert) who have several years of combined experience in conducting industrial analytics projects for industry. In the first iteration, the participants worked in 3 groups on three different use case goals with their questions and for this purpose matched the questions with the algorithm classes as described above and then collected further aspects. An excerpt from the results of one workshop is shown in figure 1.

In the first task, it was noticeable that the participants made some assumptions in order to be able to make an assignment to an algorithm class or they formulated questions as to whether certain prerequisites were given. In the second question, the participants seemed a bit lost because a starting point was missing. After some time, three main categories of aspects to be clarified emerged when the notes were grouped: product, target information, and requirements. In the second iteration, these findings were incorporated and participants provided specific information on assumptions, questions, and reformulations for three additional objectives including questions in parallel with the assignment. They were then asked to collect aspects for the three categories that are additionally important for the successful implementation of the use cases and that may need to be clarified with domain experts.

4.3 Reflection & Learning

From the observation of the participants, the recorded workshop results and additionally asked questions (“In your opinion, did the translation work in the workshop? How did

This is a resupply of March 2023 as the template used in the publication of the original article contained errors. The content of the article has remained unaffected.

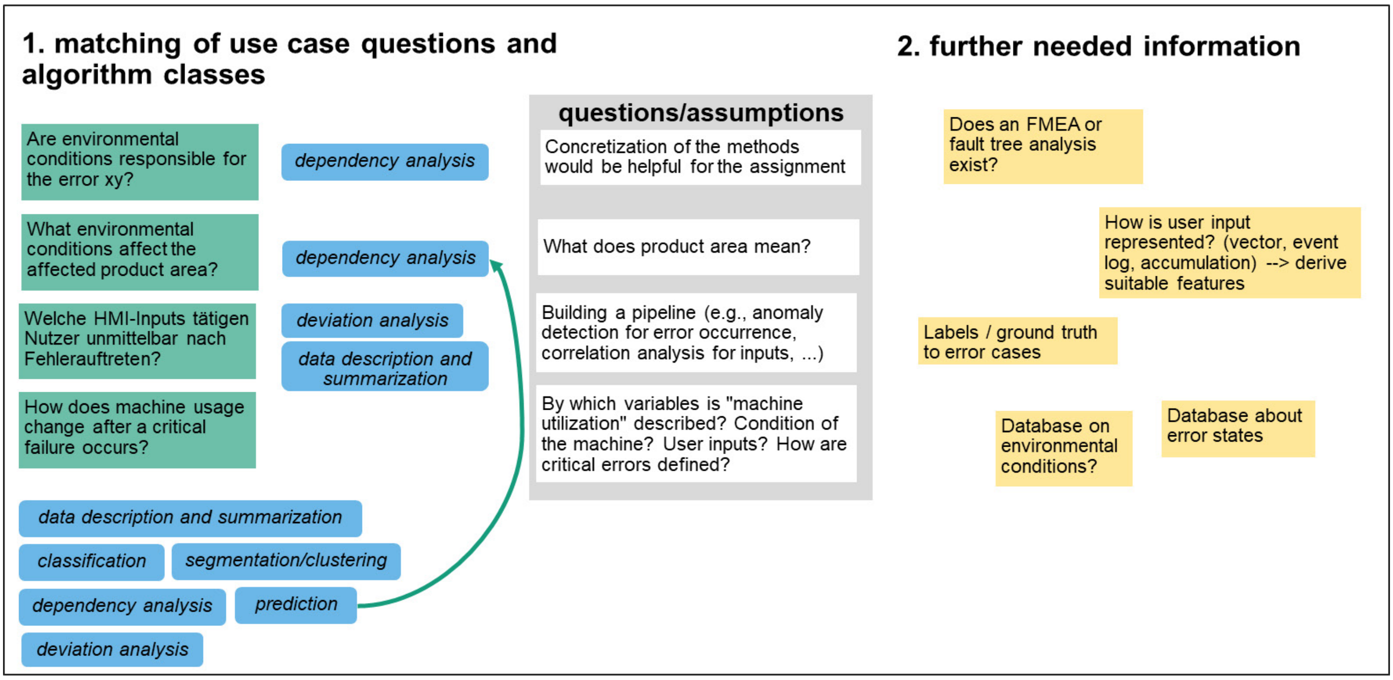

Figure 1: Excerpt from the workshop

you proceed? What problems did you encounter?"), the following learnings resulted:

1. Algorithm classes could only be assigned once the participants had answered certain questions for themselves or made certain assumptions
2. Frequent questions arose regarding the question, what individual parts of it mean exactly and by which variables these constructs are determined.
3. This results in relevant variables. Here it can be that further characteristics for the feature engineering must be derived from the available data. At this point the domain expert is again in demand.
4. It is difficult to decide on the relevant algorithm class if the structure of the data is not known.
5. In some cases, a concretization or description of the algorithm classes would help.
6. Concrete goals for each question would help
7. Questions are asked of the domain expert in the context of the product, e.g., does an FMEA or fault tree analysis exist? What is the machine usage characterized by? What settings can be made on the product?
8. Requirements arise mainly for the data, such as access to data, existence of target variables, mapping of different states in the data, uniform representation form (e.g., time stamp) and necessary data frequency.

4.4 Formulation of Learning

Derived from the experiences of the ADR project, the elements (business-)goal from the defined business use case (*learning 6*), key question(s)/hypotheses, product(context) (*learning 7*), (model-) output and factors/variables (*learning 2*) are proposed to be considered for deriving the adequate data analytics algorithm class (*learning 1*).

We converted these elements including some representative guiding questions into a shared visualization to enable collaboration within data analytics project teams on a conceptual level (see fig. 3).

At the bottom we placed the algorithm class. It can be further specified by selecting 1) the respective data analytics stage and 2) the most appropriate classes from the possible algorithm classes based on the given information. Based on existing classifications of algorithm classes or problem types of data mining [4, 14, 15] we chose the following classes: data description and summarization, classification, prediction, segmentation/clustering, dependency analysis and deviation analysis. The determination of the appropriate class can be achieved by considering the question, the goal and the desired output together in order to select the appropriate method class. In addition, algorithm class descriptions, which include typical questions or formulations, are helpful here (*learning 5*). The information element *output* can also be picked up again here and thus provide a further clue. Such a profile is shown in fig. 2.

With the fields *product* and *variables* the data scientist collects additional important information in order to understand the product and to deepen the individual variables

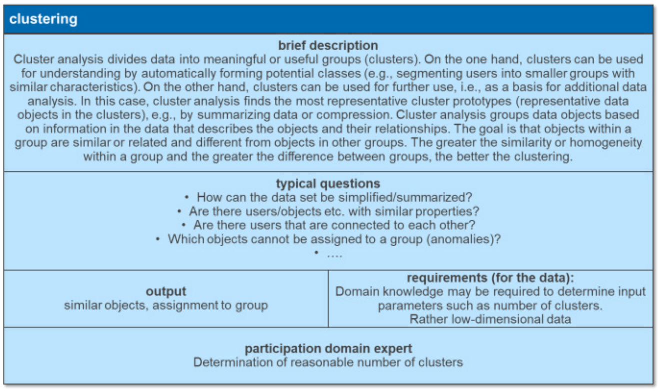

Figure 2: Example of algorithm class profile

This is a resupply of March 2023 as the template used in the publication of the original article contained errors. The content of the article has remained unaffected.

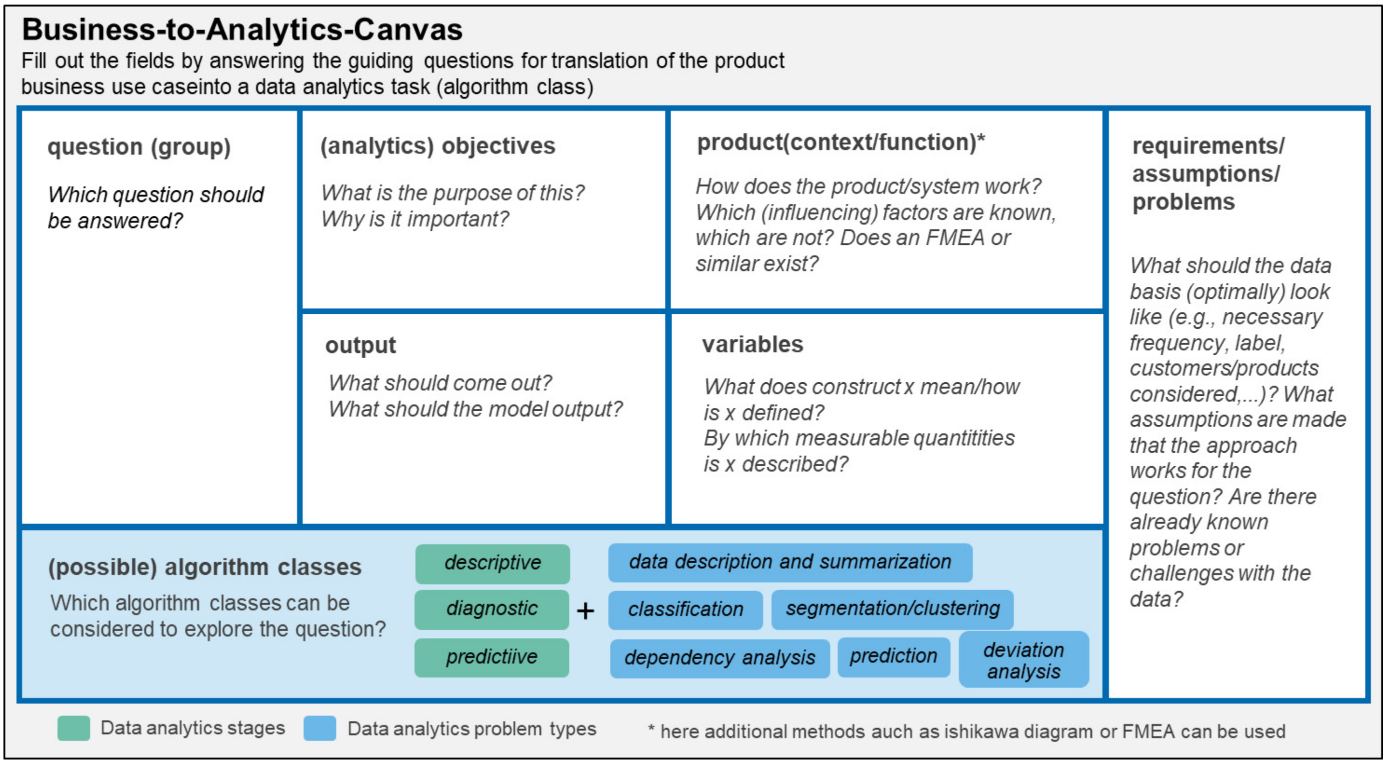

Figure 3: Business-to-analytics Canvas

or constructs of the question. A domain expert can be consulted to answer the leading questions. If necessary, it is possible to apply further methods to analyze the product in its entirety. In general, further open questions, problems and assumptions regarding the question under consideration can be collected, which should be taken into account in the further course of the data analysis (*learning 4 & 8*). The data scientist has thus conceptually specified the use case from his or her perspective.

5. Application example

The canvas was tested in a case study. The application example comes from an industrial company in the field of electrical connectivity. To optimize their product, an I/O module for machine automation, the product manager defined a use case. His business goal was to reduce service cases in the next product generation by knowing and avoiding the unstable configurations of the I/O module. It was stated that correlations between failures and module changes should be identified. In addition, particularly stable module combinations are to be identified or defined during operation: Which configurations or combinations appear together in the log files particularly often? With the help of the canvas, the use case could be transformed into a data analytics task as follows:

1. *(key) question:* Which module configurations lead to particularly frequent errors at the customer?
2. *objectives:* Better clarify the "why" to fix errors or develop new features
3. *output:* Often used configurations in connection with occurred errors
4. *product:* Modules perform different functions, such as the I/O link module can use sensors, while a module with electricity measuring function measures electric energy. Failures or errors are system crashes or a voltage drop. The assumption that modules have an influence is supported by the fact that there are modules that may only occur three times. Further influencing variables could be voltage, software (errors), parameterizations.
5. *variables:*
    a. *Error case*: e.g., voltage drop. Corresponds to error code type "error"; different codes for error cases.
    b. *Module configuration*: Module numbers
6. *requirements/assumptions/problems:*
    a. Historical service cases of several customers (logfiles in case of error) are not stored
    b. "normal" histories without error cases are not stored
    c. Information can be collected in any frequency (max. every second)

This resulted in the algorithm classes data description and dependency analysis, i.e., the analytics task is to describe trends and relationships in the data with a focus on the module configurations and to determine correlations between error cases and configurations. With this and the captured requirements and problems, the data analyst could plan the next steps. This includes, for example, the exploration and description of the existing data, the collection and storing of the missing data and the selection of a suitable algorithm of the defined algorithm classes. With the help of the canvas, the step of deriving the analytics task could be done more transparently for the project stakeholders

This is a resupply of March 2023 as the template used in the publication of the original article contained errors. The content of the article has remained unaffected.

and more quickly which contributes to the successful derivation of product optimization potential.

6. Conclusion and future work

We presented the business-to-analytics canvas, a methodological tool for transforming a product related business use case into a data analytics task on a conceptual level. It was developed using an ADR approach. Using various information elements and guiding questions, the canvas helps data scientists translate the business goal into a data analytics approach, i.e., an algorithm class, and gather the necessary information to start processing data. By defining one or more algorithm classes by application, the analytics workflow can now be designed with concrete algorithms and techniques. The canvas and its use were tested with an application example in which a communication module is to be optimized. In order to further evaluate the canvas in terms of its usefulness and usability, a user survey needs to be conducted.

Acknowledgements

This work is funded by the German Federal Ministry of Education and Research (BMBF).

References

[1] M. E. Porter and J. E. Heppelmann, “How Smart, Connected Products Are Transforming Companies (Harvard Business Review),” Online available https://hbr. org/2015/10/how-smart-connected-products are-transforming-companies. pdf, 2015

[2] M. Meyer, I. Wiederkehr, C. Koldewey, and R. Dumitrescu, “Understanding Usage Data-Driven Product Planning: A systematic literature review,” Proceedings of the Design Society, vol. 1, 2021; pp. 3289–3298.

[3] C. Shearer, “The CRISP-DM model: the new blueprint for data mining,” Journal of data warehousing, vol. 5, no. 4, 2000; pp. 13–22, 2000.

[4] P. Chapman et al., “CRISP-DM 1.0: Step-by-step data mining guide,” in 2000.

[5] S. Nalchigar and E. Yu, “Business-driven data analytics: A conceptual modeling framework,” Data & Knowledge Engineering, vol. 117, 2018.

[6] M. Meyer, I. Wiederkehr, M. Panzner, C. Koldewey, and R. Dumitrescu, “A Reference Process Model for Usage Data-Driven Product Planning,” in 2022.

[7] O. Marban, J. Gallardo, G. Mariscal, and J. Segovi, “From the Business Decision Modeling to the Use Case Modeling in Data Mining Projects,” in New Fundamental Technologies in Data Mining, K. Funatsu, Ed.: InTech, 2011.

[8] J. Segovia and O. Marbán, “Extending UML for Modeling Data Mining Projects (DM-UML),” Journal of Information Technology & Software Engineering, vol. 3, 2013.

[9] A. Kühn, R. Joppen, F. Reinhart, D. Röltgen, S. von Enzberg, and R. Dumitrescu, “Analytics Canvas – A Framework for the Design and Specification of Data Analytics Projects,” Procedia CIRP, vol. 70, 2018; pp. 162–167.

[10] Sein, Henfridsson, Purao, Rossi, and Lindgren, “Action Design Research,” MIS Quarterly, vol. 35, no. 1, 2011; p. 37.

[11] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, “A design science research methodology for information systems research,” Journal of management information systems, vol. 24, no. 3, 2007; pp. 45–77.

[12] B. Engelking, W. Buchholz, and F. Köhne, “Design principles for the application of machine learning in supply chain risk management: an action design research approach,” in Supply Management Research: Aktuelle Forschungsergebnisse 2020, C. Bode, R. Bogaschewsky, M. Eßig, R. Lasch, and W. Stölzle, Eds., Wiesbaden: Springer Fachmedien Wiesbaden, 2020; pp. 137–162.

[13] M. Meyer, M. Panzner, C. Koldewey, and R. Dumitrescu, “Towards Identifying Data Analytics Use Cases in Product Planning,” Procedia CIRP, vol. 104, 2021; pp. 1179–1184.

[14] A. Banimustafa and N. Hardy, “A Strategy for Selecting Data Mining Techniques in Metabolomics,” Methods in molecular biology (Clifton, N.J.), vol. 860, 2012; pp. 317–333.

[15] VDI/VDE 3714 Blatt 4 - Implementation and operation of big data applications in the manufacturing industry - Analysis process classes, VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik, Apr. 2021.

This is a resupply of March 2023 as the template used in the publication of the original article contained errors. The content of the article has remained unaffected.