



HAL
open science

Knowledge-Driven Data Provision to Enhance Smart Manufacturing – A Case Study in Swedish Manufacturing SME

Wei Min Wang, Helena Ebel, Steffen Kohler, Rainer Stark

► **To cite this version:**

Wei Min Wang, Helena Ebel, Steffen Kohler, Rainer Stark. Knowledge-Driven Data Provision to Enhance Smart Manufacturing – A Case Study in Swedish Manufacturing SME. 23th Working Conference on Virtual Enterprises (PRO-VE), Sep 2022, Lisbon, Portugal. pp.18-30, 10.1007/978-3-031-14844-6_2 . hal-04642011

HAL Id: hal-04642011

<https://inria.hal.science/hal-04642011v1>

Submitted on 9 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Knowledge-driven Data Provision to Enhance Smart Manufacturing – A Case Study in Swedish Manufacturing SME

Wei Min Wang¹, Helena Ebel¹, Steffen Kohler², Rainer Stark¹

¹ Technische Universität Berlin, Department of Industrial Information Technology,
Pascalstrasse 8-9, 10587 Berlin, Germany

{w.wang, helena.ebel, rainer.stark}@tu-berlin.de

² Nytt AG, Östermalmstorg 1, 114 42 Stockholm, Sweden
steffen_kohler@gmx.de

Abstract. Various novel and data-driven business concepts have emerged during the fourth industrial revolution. Smart manufacturing, for example, utilizes data from manufacturing equipment, human operators, and organizational IT systems to enable dynamic adaptations in production systems. Nowadays, these data are often distributed among multiple partners in collaborative value creation networks. Hence, to identify and collect relevant data for given business cases has become an important, but complex issue. To support the process of establishing comprehensive data provision in industrial practice, a reference model for knowledge-driven data provision processes was developed. It describes a systematic approach to drive operationalization of data provision from knowledge requirements to identify, extract and provide raw data until the application of such data sets. To evaluate the applicability of the reference model, a case study was conducted in which it was used to guide the implementation of an IoT Solution in four Swedish manufacturing companies.

Keywords: Knowledge; Data provision process; Smart Manufacturing.

1 Introduction

During the fourth industrial revolution, also known as industry 4.0, data has become the essential resource to drive innovations in the industry. Smart manufacturing (SM), as one of these novel data-driven concepts, utilizes data from manufacturing equipment, human operators, and organizational IT systems to enable integrated production systems to adapt to changes in external demands (e.g., supply networks or customer needs) in real-time [1]. Nowadays, companies often engaged in collaborative networks comprising partners with complementary expertise to provide competitive and innovative products and services. These partners, in turn, are characterized by their processes, equipment, and IT systems. In order to make data usable in such networks, the relevant data and data sources must be identified, and their transfer, processing, and subsequent application must also be carefully planned.

To implement data-driven approaches such as SM, data provision processes must therefore be considered continuous processes and regarded in the context of the strategic structuring of business processes. Respectively, the goals of implementation projects have to be defined at the very beginning starting from the intended target state after a successful implementation. From there, the required knowledge for reaching this state must be derived, and subsequently, the information required to create that knowledge. At last, the necessary data on which the information is built can be identified and collected. Various models in Knowledge Discovery and Data Mining (KDDM), such as CRISP-DM, SEMMA, and KDD, describe detailed steps to elaborate data requirements for business cases. However, all these models focus on single data mining endeavors and do not support the strategic planning and integration of continuous data provision processes as part of the organization processes.

Hence, Wang et al. proposed a reference model for knowledge-driven data provision processes (referred to as the “reference model” in the further content of this article) covering this approach [2]. By establishing a continuous data provision process, companies can apply advanced analytics and Artificial Intelligence (AI) capabilities to improve data sets’ identification, preparation, and delivery. According to the reference model (see Fig. 1), the process is initiated by business problems emerging from the companies’ value creation process. These problems lead to knowledge that needs to be fulfilled to resolve them. The satisfaction of these needs demands raw data acquired through a streamlined and automated data provision process. After the data extraction, the raw data must be transformed and applied within exploration models to obtain the information needed to satisfy the required knowledge [2].

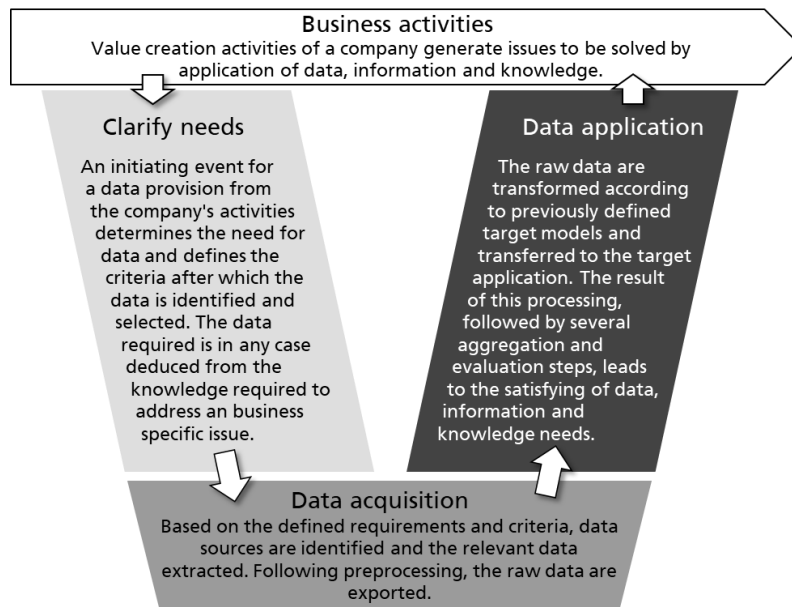


Fig. 1: The three generic phases of the reference model proposed by [2]

However, the reference model in its current state can only deliver a detailed theoretical concept for integrating knowledge-driven data provision processes. It must be examined in practical use cases to evaluate its applicability in industrial practice and identify potential improvements. Therefore, the reference model was applied in a use case that aimed to introduce SM solutions in small manufacturing companies in Sweden. Integrating a continuous data supply process is essential as SM relies on constant supply and analysis of field data. Also, to exploit optimization potentials, in-depth knowledge of the manufacturing and organizational processes is necessary to determine the relevant data obtained, suitable analysis methods, and proper ways to deliver results. The SM solution applied in the use case was provided by the Swedish company Nytt which uses IoT (internet of things) devices to monitor manufacturing processes. The gathered data are aggregated, processed, and displayed in customizable dashboards, which provide managers and machine operators with general information about their production lines (e.g., load, downtimes, outages). This information is the baseline for more advanced use cases and can be used to optimize the manufacturing processes or reduce downtime.

In the case study presented in this paper, the reference model was applied during the introduction of Nytt's solution in four Swedish SMEs. In this study, the focus was put on the "Clarify needs" phase of the model. In this phase, the knowledge objectives to be achieved are defined, determining the required information and data. Hence, in-depth interviews were carried out to examine the knowledge needs of managers and machine operators from the manufacturing companies beforehand to identify the required data and prepare to introduce the SM solution. The steps from the "data acquisition" and the "data application" phase were only regarded theoretically as Nytt's solution already included standard procedures for both. Still, the standard repertoire of data collected in the data tables of Nytt's solution could be used to evaluate if the identified knowledge needs can be explicitly mapped to actual data types.

The study's main goal was to evaluate to what extent the application of the reference model could support the identification of knowledge needs and the mapping of those to information carriers and factual data. Furthermore, it should be examined what obstacles can emerge while establishing a data provision process for a relatively simple application of the SM approach.

2 State of the Art

2.1 Smart Manufacturing

SM has attracted the attention of the manufacturing industries in recent years. Still, there are various definitions in the literature [3]. One definition of SM is given by Kang et al., based on The National Institute of Standards and Technology (NIST) [1]. It describes SM as "fully-integrated and collaborative manufacturing systems that respond in real-time to meet the changing demands and conditions in the factory, supply network, and customer needs." Kang et al. further extend this definition and explain SM as "manufacture-based technologies and systems that can respond to complicated and

diversified situations of the manufacturing field in real-time.” [1] In order to integrate responsive manufacturing systems, the manufacturing assets must be connected by technological devices like sensors, communication technology, computing platforms, and others. Monitoring the manufacturing assets in real-time improves the factory’s productivity, quality, delivery, and flexibility [1]. The application of the SM concept leads toward the connectivity of the physical manufacturing equipment layer with a cyber layer [3]. The connection is ensured by devices and sensors that acquire data from the manufacturing process [4].

The cyber-layer contains the intelligence of the smart system. It receives data from every connected machine [5] via the devices and sensors in the interface layer. Information is extracted from gathered data by applying data analytics at the cyber-layer. This information enables machines to monitor, compare performances, and predict future behavior [5]. The cyber-layer transmits decisions to the physical equipment to adjust the manufacturing process based on the gathered data. According to [3], each manufacturing equipment has intelligence [3]. This contributes to the ability of the manufacturing equipment “to cope with a complex environment and customized products” [6]. New digital concepts and technologies such as digital twins, IoT, and cloud are being established alongside data-driven models to ensure that the required intelligence is technically implemented in SM [7].

2.2 Knowledge Discovery and Data Mining

The required data and information flow must be clarified to enable a proper level of automation in manufacturing systems. Different approaches can be found to describe the flow of information in the literature. Forza and Salvador [8] classify the information flow within companies into three groups regarding their direction: vertical, horizontal, and external information flow. The international standard IEC 62264-1 [9] itself presents a functional data flow model, including the interactions between the activities at the manufacturing operations management level and business planning and logistics level. This functional model varies from the vertical structure of a company and leads to a complex outline of the vertical data flow between the above two levels. Lower levels as the actual production process are not included in the model. Tao et al. [10] present a manufacturing data life cycle involving seven components for transforming data into understandable information. The increasing application of sensors and connected machines leads to the creation of high-volume data (big data) [5]. This results in difficulty extracting relevant information and knowledge hidden within those data sets [10]. By obtaining the knowledge, industrial practitioners hope to reveal potential improvements in their organizational processes [2].

Hence, collecting and analyzing of high quantities of data has become crucial. Approaches to extract information and knowledge from large amounts of data cover Knowledge Discovery (KD), Data Mining (DM), Knowledge Discovery in Databases (KDD), or KDDM. The terms are frequently reviewed in various academic papers in the past years, and various opinions have emerged regarding the differentiation of the concepts. Some academic papers define DM only as part of KDD’s overall knowledge discovery process [11]. Others describe DM as a synonym for the KDD process [12, 13]. Lately, the term KDDM was suggested as an overall name for the process [14].

Despite differences in understanding, all approaches aim at revealing hidden information and knowledge from large amounts of data. The increasing attention for KDDM is related to techniques and methods from machine learning and AI to exploit hidden knowledge from huge data sets [15]. Still, the existing KDDM models mainly focus on single data mining projects and do not support establishing continuous data provision processes.

Another well-known procedure for the application of DM is CRISP-DM, which is based on KDD approaches [16]. CRISP-DM is a methodology to discover the knowledge of a business use case with DM [17]. Although this methodology primarily focuses on business understanding and identifying business goals for the use case, it has a gap in business processes, where several activities are done to achieve a business goal. In addition, CRISP-DM assumed that all data is known and already available [18], which is not necessarily the case with new business models and resulting business processes.

Data requirements for novel data-driven products and business models lead to continuous data provision processes that are indispensable based on the organizations' strategic orientation and enable constant data acquisition and knowledge discovery. Furthermore, a continuous knowledge-driven data provision process enables AI capabilities and advanced data analytics to exploit the collected data and improve the organizational processes [2].

3 Case Study

3.1 Case Setting

The foundation for SM is connected manufacturing systems and respective means to collect machine and process-related data. Nytt has developed a production monitoring solution based on smartphones that can be added to various types of manufacturing machines. The product is integrated into the physical production process and extracts data directly from the manufacturing operations. The primary function of the solution is to monitor the current manufacturing process and continuously collect real-time data from the ongoing operations. It enables the exchange between the physical production assets and computer-based systems through wireless networks. The product involves a system consisting of three main components: the *SetApp*, the *OpApp*, and the *dashboard* to create this connectivity. Data gathered by the *SetApp* (process data) and the *OpApp* (operator and machine data) are processed and displayed via standardized graphs within the *dashboard* to overview the manufacturing assets. These graphs display predefined production metrics, e.g., productivity or production volume, and can be customized by the user through different filters, e.g., time. However, the available data were limited to the given capabilities of Nytt's solution at that time. They included captured machine state, time of the capture, input from machine operators, and additional sensor data, i.e., input-output measuring, acceleration, brightness, and volume level.

3.2 Evaluation Goals and Methods

The goal of the use case was to evaluate two main research questions (RQ) that can shed light on the applicability of the reference model proposed by [2]:

- RQ1: Can knowledge emerging from the current business situation of the hosting companies be explicitly mapped to data that is collectible from the manufacturing process?
- RQ2: What challenges emerge from integrating a continuous data provision process in a realistic environment?

To answer RQ1, two kinds of data had to be collected: firstly, the knowledge needs of the interviewees, and secondly, the operational data from the manufacturing processes. According to [2], the definition of the knowledge needs is the foundation for target-oriented data provision processes as they allow for identifying the relevant information and data to be collected. In the current case study, the knowledge needs of the interviewees were used to identify concrete data needs and map them to available data in the data tables of Nytt's solution. If the required data was available in the data tables, the according to visualization graph was selected and presented to the interviewees asking them to determine whether their knowledge needs could be satisfied with the given data and their visualization.

As an evaluation method for the knowledge needs and their fulfillment by the presented visualization, semi-structured interviews were chosen as they promised a balanced approach to collect comparable results while allowing the interviews to be led in conversational manners to explore the thoughts and desires of the interviewees [19, 20]. In total, interviews with seven employees at five Swedish manufacturing companies (A-E) were arranged, and Nytt's product was introduced. All companies were small-sized companies. Managers (n=5) and machine operators (n=2) were interviewed to obtain insights from different perspectives.

All interviews lasted between 26 and 35 minutes and were conducted through video calls. The interviews were all held in English. Following [20], all interviews were recorded and transcribed after they were conducted. These transcripts were the basis for the interview analysis. The interview results were analyzed using qualitative content analysis, namely the ordinal deductive category assignment method [21].

4 Results

4.1 Results Regarding RQ1

During the interviews, a total of 21 knowledge needs could be identified in the four companies (see Table 1). 13 of these knowledge needs could be directly mapped to concrete data collected by Nytt's solution. The required data could be identified for the eight remaining knowledge needs, but they were unavailable as they were not or only partially collected by Nytt's solution at that time.

Table 1: Identified knowledge needs and availability of concrete data

Nr.	Knowledge need	Com-pany	Could be mapped to available data? (yes / no)
KN1	How are the machines currently performing, especially how much output do they currently generate?	A	yes
KN2	When did the performance in the past not meet the expectations, and why?	A	yes
KN3	How much time is necessary to prepare the machines?	A	yes
KN4	What is the exact time the single manufacturing jobs require, and how much interference is necessary?	A	no
KN5	How can I obtain the output of the production process at any requested time without performing a manual data collection process	B	yes
KN6	What number of production cycles have the machines been running with each worn-out tool?	B	yes
KN7	How much of the production time was the equipment not producing?	B	yes
KN8	How are the machines on the shop floor currently performing, and what is the current uptime of the machines?	B	yes
KN9	How much time do we spent on reworking defective parts?	B	no
KN10	How can I overview the key metrics and stops of my manufacturing equipment at all times without performing a manual data collection process?	C	yes
KN11	Does the performance of the current production operations meets the re-quired level? Did it exceed this level, or did it fall behind?	C	yes
KN12	What is the status of the assigned maintenance tasks at any time?	C	yes
KN13	How can I overview the performance and stops of my manufacturing machines at any time?	D	yes

KN14	How much time do the employees need to assemble a product?	D	yes
KN15	What is the necessary information a new employee needs, to be able to operate the machines after a shorter teach-in phase?	D	no
KN16	What is the output of parts of the production process at any requested time?	D	no
KN17	What kind of maintenance service do the manufacturing equipment require and when?	D	no
KN18	How can I collect data from the older machines to monitor their performance without a costly and complex installation process of sensors?	E	yes
KN19	How high is the strength of the Wi-Fi signal, the brightness, or sound volume in the factory at all time?	E	no
KN20	What is the filling level of the pallets in the production's lines, and how can I receive automated notifications if the filling level is low?	E	no
KN21	How can I locate AGVs, employees on the shop floor, and products in real-time?	E	no

Eleven graphs from the portfolio currently provided by Nytt's solution could be identified that would satisfy these needs (see Table 2). The following section describes how these graphs were used to satisfy the knowledge needs.

Table 2: Graph types that supports the satisfaction of knowledge need

Nr	Graph type
1	Productivity and the "production", "setup", and "no production" time
2	Daily Overview
3	Production Volume
4	Throughput Rate
5	Total number of Stops (categorised in four classes according to length of the stops)
6	Scatter plot of Stop distribution with complementary display of reasons for stops (as bar chart)
7	Maintenance tasks
8	Assembly activities

9	Development production metrics over time
10	Production cycles
11	Comparing Production Volume

KN1 - “How are the machines currently performing, especially how much output do they currently generate?” This knowledge need is answered by illustrating the metrics used in company A to measure manufacturing performance. These metrics are “Productivity”, “Production Volume”, and “Production Time”. To evaluate the current performance, real-time data from the manufacturing processes are transformed into these metrics. Thus, the metrics continuously reflect the current manufacturing performance. The knowledge need can be satisfied with graphs 1, 2, and 3. Satisfying the knowledge need also significantly improved the manufacturing performance of a specific machine with an automated loading mechanism and produced almost independently. The autonomous loading mechanism and the machine's good condition result in higher productivity than 90 percent. Due to the high productivity, the managers/production managers could not increase the output during the regular shift. By evaluating the visualized production data over several weeks, the most significant potential to increase the output was the exploitation of the production time between midnight and 7 am. The production output increased by almost 15 percent by satisfying the urge for higher transparency. The satisfaction of a knowledge need did unlock potential improvement of the organizational processes at the use case. It proves that business problems can be solved and value creation improved, as suggested by [2].

KN2 - “When did the performance in the past not meet the expectations, and why?” Fulfilling this knowledge need leads to capturing the performance over time and displaying it in a manner that the user has the chance to compare past performances. Graph type 9 illustrates the development of the metrics over time, can consequently satisfy the knowledge need, and allows company A to conclude past performances.

KN3 - “How much time is necessary for preparing the machines?” This question can be satisfied by displaying the machine's overall "setup" time, which is presented daily, weekly, monthly, or yearly in graph 1.

KN5 - “What is the output of the production process at any requested time without performing a manual data collection process?” KN5 can be answered by graph type 3 as it provides the production volume in real-time. Furthermore, applying filters enables examining the production volume of past days, weeks, months, and years. The satisfaction of this knowledge needs to lead to an increased production volume by 6 percent of company B. However, unlike company A, not the graph enabled the improvement, but the “Raw Data.”

KN6 - “What number of production cycles have the machines been running with each worn-out tool?” The knowledge need is satisfied by graph type 10. It con-

tains the number of production cycles that the machines conduct. Additionally, by selecting a specific time, company B can look up the number of production cycles before every tool change.

KN7 - “How much of the production time was the equipment not producing?”

The illustration of the "no production" time and "setup" time in graph type 1 satisfies the required knowledge. The two figures can be displayed daily, weekly, monthly, or yearly.

KN8 - “How are the machines on the shop floor currently performing, and what is the current uptime of the machines?” Similar to company A, company B measures its performance with manufacturing metrics. These metrics are “Production Volume”, “Production Time”, and “Equipment Downtime”. The information about these metrics is displayed in graphs 1, 2, and 3. The real-time visualization creates the desired overview of the performance. The machine uptimes can be displayed in a time unit (hours) or as a percentage share of the overall "production" time.

KN10 - “How can I overview the key metrics and stops of my manufacturing equipment at all times without performing a manual data collection process?”

Company C measures its performance with the metrics “Productivity”, “Production Volume”, and “Production Time”. Graph types 1, 2, and 3 provide real-time information reflecting these metrics to satisfy the knowledge need. Furthermore, the graph types 5, and 6 illustrate each machine stop alongside the metrics.

KN11 - “Does the performance of the current production operations meets the required level? Did it exceed this level, or did it fall behind?” The knowledge need is answered by comparing the current performance metrics to values the metrics should reach by the end of the preferred period. The interviewee named the production output as a key metric for measuring the performance level. Graph type 11 provides the information required to satisfy the knowledge need.

KN12 - “What is the status of the assigned maintenance tasks at any time?”

Graph type 7 and the detailed table are essential to answer this knowledge need. The information shows which and how many tasks are completed.

KN13 - “How can I overview the performance and stops of my manufacturing machines at any time?” This knowledge need is equivalent to KN10 of company C. Thus, it can be satisfied similarly. The relevant metrics to measure the manufacturing performance at company D are “Production Volume” and “Throughput Rate”. This information is illustrated by the graph types 3, 4, 5, and 6.

KN14 - “How much time do the employees need to assemble a product?” This knowledge need is closely related to the production metric “Through-put Rate”. The interviewee from company D wished to understand the duration of the assembly activities better. Graph type 8 shows the duration of the single assembly activities and the preparation time of these activities. This graph satisfies the knowledge need.

KN18 - “How can I collect data from the older machines to monitor their performance without a costly and complex installation process of sensors?” Company E is not applying NYTT’s product in their production processes but as a tool for proof of concept. Therefore, the interviewee from company E did not request any specific production metrics or figures. The currently available graphs in the dashboard were sufficient to conduct a limited amount of tests. The information is available in graphs types 1, 2, and 3.

4.2 Results Regarding RQ2

Interviewing employees from two different groups at the five companies did not only lead to further insights into the use cases themselves but also revealed different perceptions of the benefits provided by the continuous data provision process. At company A and company D, two employees from different positions participated in the interviews. The answers included significant variations in the expressed knowledge needs and acceptance of the continuous data provision process. While the two interviewees from the manager/production manager position are relatively positive toward the data provision process, the two machine operators agree that the process does not benefit their daily tasks. The two managers/production managers seem more curious about the possible benefits the product can provide them. They were expressed several knowledge needs that could be satisfied applying of the continuous data provision process.

Furthermore, the two managers/production managers are the managing directors of their respective companies, and they initiated the use of the data provision process. That indicates that the use of the product aligns with the organizational goals they are following. The answers given by the two managers/production managers can be reflected in the statements from the other interviewed managers/production managers. The managers/production managers demanded data from their manufacturing processes and expressed similar knowledge needs. That confirms the favorable position of the managers/production managers towards applying a continuous data provision process. The opinions of the two machine operators from different companies showed the same analogy, but they did not see the benefits of the data provision process.

5 Discussion

The analysis of the interview results showed that all interviewees could conclude the detailed data needed for the knowledge they stated before. This shows that the knowledge needs can be captured explicit enough through a semi-structured interview to be mapped to collectible data, as suggested by [2]. The information provided by the graphs enabled the identification of improvement potentials in all companies. In two cases, the identified potentials led to direct actions. This supports the hypothesis by [2] that the reference model can support establishing data provision processes that improve a company’s value creation process. Consequently, RQ1 could be answered positively.

This evaluation showed that the company size significantly influenced continuous data provision processes in realistic environments. It was observed that the integration of data provision processes was easier in smaller companies due to less complex value creation processes, short and direct communications, and more significant influences of single employees. Moreover, in smaller companies, the effort to involve the entirety of the employees was lower, leading to a higher level of staff commitment. For RQ2, it can be stated that – at least in this sample – the challenges did not emerge from data-related aspects (e.g., availability, accessibility, quality) but from organizational aspects, mainly company size and the commitment of the personnel.

6 Conclusion and Outlook

The current case study showed that the reference model proposed by [2] could be applied in an industrial context. As the results show, the process of integrating Nytt's IoT solution for smart manufacturing could be supported by the model, independently of the size of the company, their process complexity, or the level of digitization. The data collected could be mapped to the identified knowledge needs in most cases. Thus, the data could be integrated into a continuous process to supply specialized graphs in the dashboards. Using these dashboards, all companies in the study sample could identify improvement potentials or even take direct actions.

However, the case study revealed some significant limitations to the reference model. One limitation relates to the strong foundation of the reference model on the knowledge needs of the individuals applying the model. Individuals usually try to solve problems based on their experiences and routines, but the knowledge that needs to be derived is limited to the individuals' imagination of the solution. That can lead to a lack in exploiting the full use of a continuous data provision process, especially for companies with a low level of digitalization, as problems are solved without exploiting modern tools. The second limitation is based on the different perceptions of the usefulness continuous data provision process during this study. Some interviewees shared a prevail of manual work over the benefits of the process. They questioned the value of a data provision process as they were worried about getting additional work without immediate benefits. That applies especially to the individuals responsible for maintaining the continuous data provision process through manual work. For a successful application of the reference model in realistic environments, the benefits and efforts of every individual involved should be considered and communicated transparently.

Furthermore, the results of this case study are also limited by its setting. As this case study only included small-sized companies, the reference model will need further evaluations in the context of bigger companies with more complex organizational structures. Also, the "data collection" and "data application" phase of the reference model could only be considered in a theoretical manner as Nytt's solution already included standardized procedures for both phases. Consequently, both aspects must be further investigated in academic research or industrial use cases.

References

1. Kang, H.S., Lee, J.Y., Choi, S., Kim, H., Park, J.H., Son, J.Y., Kim, B.H., Noh, S.D.: Smart manufacturing: Past research, present findings, and future directions. *International Journal of Precision Engineering and Manufacturing-Green Technology*, vol. 3, 111–128 (2016)
2. Wang, W.M., Preidel, M., Fachbach, B., Stark, R.: Towards a Reference Model for Knowledge Driven Data Provision Processes. In: Camarinha-Matos, L.M., Afsarmanesh, H., Ortiz, A. (eds.) *Boosting Collaborative Networks 4.0*, pp. 123–132. Springer International Publishing (2020)
3. Kusiak, A.: Smart manufacturing. *International Journal of Production Research*, vol. 56, 508–517 (2018)
4. Jiang, J.-R.: An improved cyber-physical systems architecture for Industry 4.0 smart factories. *Advances in Mechanical Engineering*, vol. 10, 15 (2018)
5. Lee, J., Bagheri, B., Kao, H.-A.: A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, vol. 3, 18–23 (2015). doi: 10.1016/j.mfglet.2014.12.001
6. Brettel, M., Friederichsen, N., Keller, M., Rosenberg, M.: How Virtualization, Decentralization and Network Building Change the Manufacturing Landscape: An Industry 4.0 Perspective. *World Academy of Science, Engineering and Technology, International Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering*, vol. 8, 37–44 (2014)
7. Stark, R.: *Virtual Product Creation in Industry*. Springer Berlin Heidelberg, Berlin, Heidelberg (2022)
8. Forza, C., Salvador, F.: Information flows for high-performance manufacturing. *International Journal of Production Economics*, vol. 70, 21–36 (2001)
9. ISO: IEC 62264-1: 2013. *Enterprise-control system integration — Part 1: Models and terminology*
10. Tao, F., Qi, Q., Liu, A., Kusiak, A.: Data-driven smart manufacturing. *Journal of Manufacturing Systems*, vol. 48, 157–169 (2018)
11. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, vol. 17, 37–54 (1996)
12. Cios, K.J., Kurgan, L.A.: Trends in Data Mining and Knowledge Discovery. In: Pal, N.R., Jain, L. (eds.) *Advanced Techniques in Knowledge Discovery and Data Mining*, pp. 1–26. Springer London, London (2005)
13. Han, J., Kamber, M., Pei, J.: *Data mining. Concepts and techniques*. Morgan Kaufmann/Elsevier, Waltham, MA (2012)
14. Mariscal, G., Marbán, Ó., Fernández, C.: A survey of data mining and knowledge discovery process models and methodologies. *The Knowledge Engineering Review*, vol. 25, 137–166 (2010)
15. Smyth, P.: Data mining: data analysis on a grand scale? *Statistical Methods in Medical Research*, vol. 9 (2000)
16. Azevedo, A., Santos, M.F.: KDD, SEMMA and CRISP-DM: A Parallel Overview. *IADIS European Conference on Data Mining*, Amsterdam, The Netherlands, vol. (2008)
17. Wirth, R., Hipp, J.: CRISP-DM: Towards a Standard Process Model for Data Mining. *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining.*, vol. 1, 29–40 (2000)
18. Wiemer, H., Drowatzky, L., Ihlenfeldt, S.: Data Mining Methodology for Engineering Applications (DMME)—A Holistic Extension to the CRISP-DM Model. *Applied Sciences*, vol. 9, 2407 (2019). doi: 10.3390/app9122407

19. Loosen, W.: Das Leitfadeninterview – eine unterschätzte Methode. In: Averbek-Lietz, S., Meyen, M (ed.) Handbuch nicht standardisierte Methoden in der Kommunikationswissenschaft, pp. 139–155. Springer VS, Wiesbaden (2016). doi: 10.1007/978-3-658-01656-2_9
20. Longhurst, R.: Semi-structured interviews and focus groups. Key methods in geography, vol. 3, 143–156 (2003)
21. Mayring, P.: Qualitative content analysis: theoretical foundation, basic procedures and software solution, vol. (2014)