Enabling Manufacturing Companies to Implement Data-based Applications to Support Lean Practices

DISSERTATION
of the University of St. Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Management

Submitted by

Paul Bueß
from
Germany

Approved on the application of
Prof. Dr. Thomas Friedli
and
Prof. Dr. Kuno Schedler

Dissertation no. 4983

Difo-Druck GmbH, Untersiemau 2020

The University of St.Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, May 18, 2020

The President:

Prof. Dr. Bernhard Ehrenzeller

Vorwort des Autors

Die vorliegende Dissertation entstand im Rahmen meiner Tätigkeit als wissenschaftlicher Mitarbeiter im Bereich Produktionsmanagement am Institut für Technologiemanagement der Universität St. Gallen. Meine Rolle im Team für Operational Excellence erlaubte es mir früh, die Thematik der schlanken Produktion aus einer theoretischen und praktischen Perspektive zu diskutieren. Im Rahmen eines Benchmarking Projektes sowie eines KTI Projektes rückte zunehmend auch das Thema Industrie 4.0 in mein Blickfeld. Die Erfahrungen aus diesen Projekten und insbesondere der intensive Austausch mit verschiedenen Forschungspartnern dienen als Grundlage für diese Arbeit. Daher möchte ich einen besonderen Dank an alle beteiligten Partnerunternehmen für die Offenheit zum Austausch aussprechen. Ausserdem möchte ich die Bereitschaft von Professor Schuster sowie Professor Wuest zur Unterstützung der Arbeit herausheben. Beide standen trotz voller Kalender einer Diskussion sowie einem Experteninterview von Anfang an offen gegenüber.

Ganz besonderer Dank gilt meinem Doktorvater Professor Thomas Friedli. Ohne seine Unterstützung hätte ich ein Thema an der Schnittstelle von Lean Manufacturing und Industrie 4.0 nicht bearbeiten können. Er ermöglichte es mir, mich auch in Projekten anderer Gruppen zu involvieren, um so die nötige Empirie für diese Arbeit zu sammeln. Er liess mir dabei die notwendigen Freiheiten in der Ausrichtung und Bearbeitung des Themas, stand jedoch stets bei Fragen zur Verfügung und lieferte wichtige Impulse für diese Arbeit. Ebenso möchte ich mich bei Professor Kuno Schedler für die Übernahme des Koreferats bedanken.

Ein herzliches Dankeschön möchte ich an alle meine Kollegen richten, mit denen ich nicht nur erfolgreich Projekte bearbeitet habe, sondern auch viele interessante Diskussionen geführt und großartige Momente im Büro, beim Töggeli oder in den Bergen erleben durfte.

Die Möglichkeit zur Promotion verdanke ich zu erheblichen Teilen meinen Eltern und meinen Großeltern. Sie sind stets mit viel Geduld meiner Neugier begegnet und haben mich immer bei meinen Vorhaben unterstützt. Meine Großeltern, bei denen ich in der Kindheit viel Zeit verbracht habe, haben mir auf eindrucksvolle Weise vermittelt, dass es sich lohnt für ein Ziel zu arbeiten, auch wenn es dabei Widerstände zu überwinden gibt.

Zuletzt möchte ich meiner Freundin Ines dafür danken, dass sie mich während der gesamten Promotionszeit unterstützt hat und Verständnis aufbrachte, wenn die Arbeit viel von der knappen gemeinsamen Zeit in Anspruch nahm. Sie zeigte mir immer wieder, dass es ein Leben neben der Dissertation gibt und sorgte damit dafür, dass ich die Dissertationszeit insgesamt als sehr positiv in Erinnerung habe.

St. Gallen, Oktober 2019 Paul Bueß

Summary

The manufacturing industry is still a critical sector to generate economic wealth in many countries. From the end of the 18th century, the manufacturing industry has undergone several paradigm shifts from craftmanship to western-dominated mass production and eventually to lean manufacturing (LM). LM builds on the Toyota Production System and has enabled companies around the globe to increase quality and productivity by eliminating waste. Nevertheless, companies are pressured to increase the performance of their lean production system (LPS) to meet steadily increasing customer demands. Literature suggests that LM may benefit from the integration of emerging smart manufacturing (SM) technology. A survey on the next stage of lean indicates that especially data collection and analysis will be a key driver to increase the performance of LPSs. Existing literature on the integration of SM into LM, however, tends to stay on a generic level and fails to address the impact of SM technologies on concrete lean practices.

These observations motivated the research of the question of how manufacturing companies can be enabled to implement data-based applications (DBAs) to support lean practices. To answer this question three sub-research questions (SRQ) have been derived. SRQ 1 addresses the question: "which DBAs exist and what are their objectives" by conducting a comprehensive literature review and providing a structured classification of DBAs in manufacturing. SRQ 2 documents challenges and enablers for DBAs based on qualitative studies, including case studies with awarded successful practice companies and expert interviews with two senior academics in the field of data utilization in manufacturing. Finally, SRQ 3 bridges the gap of data utilization and LM by evaluating the potential of DBAs to support 10 established lean practices. The methodology follows a pairwise DBA—lean practice impact evaluation, resulting in the DBA—Lean Practice Impact Matrix.

Accordingly, the structured DBA overview and the DBA—Lean Practice Impact Matrix—are key contributions of this dissertation. The former allows manufacturing managers to identify opportunities to capitalize on their existing manufacturing data, while the latter provides impulses for lean managers on how to improve the performance of their LPS by using DBAs. The documentation of challenges and enablers allows readers to learn how real-world challenges can be addressed and thus enable inter-organizational learning. In addition, the research proposes a theoretically backed approach to better understand an observed hesitation of manufacturing companies to invest in DBA, although high potential is expected.

In conclusion, in the broader context of the interplay of LM and SM, this dissertation can serve as a starting point to better understand the potential but also the challenges of exploiting manufacturing data to enable LPSs to meet steadily increasing customer demands.

Zusammenfassung

Die produzierende Industrie ist nach wie vor ein wichtiger Sektor für den Wohlstand in vielen Ländern. Sie hat seit dem Ende des 18. Jahrhunderts mehrere Paradigmenwechsel durchlaufen, vom Handwerk über die Massenproduktion bis hin zur schlanken Produktion, Lean Manufacturing (LM) gennant. LM ermöglichte es Unternehmen, Qualität und Produktivität durch Vermeidung von Verschwendung zu steigern. Stetig steigenden Kundenanforderungen verlangen jedoch von Unternehmen die Leistungsfähigkeit ihrer Lean Produktions Systeme (LPS) permanent zu erhöhen. Die Literatur legt nahe, dass LM von der Integration neuer Industrie 4.0 Technologie profitieren kann. Eine Umfrage zur nächsten Entwicklungsstufe von Lean zeigt zudem, dass besonders die Datennutzung ein wichtiger Faktor zur Steigerung der Leistung von LPSs sein wird. Bestehende Literatur über die Integration neuer Technologien in LM bleibt jedoch oft auf einer allgemeinen Ebene, sodass eine Analyse der Effekte auf konkrete Lean-Praktiken fehlt.

Diese Beobachtungen begründen die Fragestellung, wie Produktionsunternehmen in die Lage versetzt werden können, datenbasierte Anwendungen (DBAs) zur Unterstützung von Lean-Praktiken zu implementieren. Um diese Frage zu beantworten, wurden drei Unterforschungsfragen (UFF) abgeleitet. UFF 1 befasst sich mit der Frage: "welche DBAs gibt es und was sind ihre Ziele" und beantwortet diese mithilfe einer umfassende Literaturrecherche, die eine strukturierte Klassifizierung bestehender DBAs ermöglicht. UFF 2 dokumentiert Herausforderungen und Befähiger für die Anwendung von DBAs, welche sich aus Fallstudien mit ausgezeichneten Praxisunternehmen sowie zwei Experteninterviews ableiten. Schließlich schließt UFF 3 die Lücke zwischen Datennutzung und LM, indem es das Potenzial von DBAs zur Unterstützung von zehn etablierten Lean-Praktiken bewertet. Die Methodik folgt einer paarweisen DBA-Leanpraktik-Bewertung, welche die Grundlage für die DBA-Lean-Praktik Einfluss Matrix darstellt.

Die strukturierte DBA-Übersicht und die DBA-Leanpraktik Einflussmatrix sind wesentliche Beiträge dieser Arbeit. Erstere ermöglicht es Managern Chancen zur Nutzung ihrer vorhandenen Produktionsdaten zu identifizieren, während letztere Lean Managern Impulse geben kann, wie sie die Leistungsfähigkeit ihres LPS durch den Einsatz von DBAs steigern können. Die Dokumentation von Herausforderungen und Befähigern vermittelt, wie reale Herausforderungen von DBA Projekten angegangen wurden. Darüber hinaus schlägt die Arbeit einen theoretischen Ansatz vor, um eine beobachtete Zurückhaltung produzierender Unternehmen in DBAs zu investieren, zu erklären. Im Kontext des Zusammenspiels von LM und I4.0 kann diese Dissertation als Ausgangspunkt dienen, um das Potenzial, als auch die Herausforderungen der Nutzung von Produktionsdaten besser zu verstehen, um so LPSs zu befähigen weiterhin steigende Kundenanforderungen zu erfüllen.

Table of Content

List of Figures

List of Tables

List of Abbreviations

AI	Artificial intelligence
AGV	Automated guided vehicle
CIM	Computer integrated manufacturing
CI	Continuous improvement
CPS	Cyber-Physical Systems
DM	Data mining
DBA	Data-based application
et al.	et alii
ERP	Enterprise resource planning system
ICT	Information and communication
ITEM-HSG	Institute for Technology Management, University of St. Gallen
IoT	Internet of things
JIT	Just-in-time
KPI	Key performance indicator
LM	Lean manufacturing
LPS	Lean production system
ML	Machine learning
MRQ	Main-research-question
MAS	Manufacturing Analytics Solution (Use case Company A)
MES	Manufacturing execution system
OM	Operations management
OEE	Overall equipment effectiveness
RFID	Radio frequency identification device
ROI	Return on investment
SM	Smart manufacturing
SOT	Standard operating time
SRQ	Sub-research-question
SP	Successful practice (company)
TAM	Technology acceptance model
TPM	Total productive maintenance
TQM	Total quality management
TPS	Toyota production system
VSM	Value stream mapping
WIP	Work in process

1 Introduction

1.1 Motivation and Relevance

1.1.1 Practical Relevance

"Data is the new oil. It's valuable, but if unrefined it cannot really be used. It has to be changed into gas, plastic, chemicals, etc. to create a valuable entity that drives profitable activity; so must data be broken down, analyzed for it to have value."

Clive Humby

In 2018 the five most valuable companies in the world, measured by their market capitalization, are IT companies. At the same time, only one manufacturing company is listed among the top 10 most expensive enterprises. However, market capitalization is only one aspect of the importance of a company. Another, arguably at least equally important aspect, is the number of jobs a company provides. Looking at this number, the major relevance of the manufacturing sector becomes apparent. The Volkswagen AG alone employs almost twice as many employees as Apple, Google, and Microsoft together[1]. Furthermore, in Germany there are more than five times as many people in the manufacturing industry as in the IT sector[2]. Manufacturing, therefore, is still a critical industry for generating economic growth and economic wealth (Y. Chen, 2017, p. 588).

The history of manufacturing has shown several paradigm shifts. Around 1900, the production of goods was dominated by craftsmanship. In 1908, Henry Ford introduced the Model T and in 1913 the first moving assembly line. This was the starting point of mass production, which was the dominating production paradigm until the emergence of lean manufacturing (LM) after World War II with its origins in Japan. Thanks to lean, Japanese companies achieved significant competitive advantages over their U.S. competitors in the following years. Around 1990, Toyota, which created the blueprint for LM, was considered the most efficient and highest-quality car manufacturer worldwide (Womack, Jones, & Roos, 1990, p. 49). Since then, Toyota stayed highly successful and is today the second-largest automotive company worldwide[3]. LM has enabled many companies worldwide to increase their productivity and quality by eliminating waste. Holweg (2007, p. 420), therefore, characterizes LM as the current most influential manufacturing paradigm.

---

1 Number of full-time employees as of end of 2017 (in 1000): Apple (2018) : 123, Microsoft (2018) 114, Alphabet (2018) : 88, Volkswagen (2018) : 642

2 Number of full-time employees as of end of 2017 (in 1,000):
Manufacturing sector: 5.300 (Statistisches Bundesamt (2018)), IT – sector 950 (Bitkom (2018))

3 Based on cars sold in 2017, not considering the alliance of Renault – Nissan –Mitsubishi (Statista (2018))

However, steadily increasing customer demand for highest quality, cost efficiency, and product variability (Gerberich, 2011, p. 392) as well as global competition, pressures manufacturers to further increase the effectiveness and efficiency of their lean production system (LPS).

In a cross-industry benchmarking study on lean production in 2017 (Macuvele, Buess, & Friedli, 2018), participants reported decreasing productivity gains by standard lean tools and methods: the low hanging fruits in lean production have already been picked. Consequently, lean companies need to find new ways to improve their production system to meet customer demands.

Besides LM, another manufacturing concept recently received increased attention from academia, media, and government. IT technology-driven smart manufacturing (SM) is heavily supported by governmental programs in many of the most important manufacturing countries (Thoben, Wiesner, & Wuest, 2017, p. 5). Lasi, Fettke, Kemper, Feld, and Hoffmann (2014, p. 239) argue that SM technology will lead to the next fundamental paradigm shift in manufacturing. However, even though most of the participants of the 2017 lean study recognize the potential of new IT technology, more than 90 percent are convinced that SM will not replace LM. Instead, new IT technology is considered as a complementary enabler to achieve the next level of LM.

A core element of LM is continuous improvement (CI), which is the relentless search for improvement opportunities. Most CI methods rely on data to get a sound understanding of the problem and its root cause. Consequently, participants of the lean study have identified data availability and transparency, provided by state-of-the-art IT technology, as promising resources for further operational improvements. This finding is supported by Qi and Tao (2018, p. 3591) who argue that data will support the identification of visible and invisible problems in complex manufacturing processes. According to O’Donovan, Leahy, Bruton, and O’Sullivan (2015b, p. 2), data-driven preventive maintenance has a positive effect on operating costs with a saving potential of more than 30 percent. Harding, Shahbaz, Srinivas, and Kusiak (2006, p. 969) identified the manufacturing sector as an industry where data mining can contribute significantly to companies' competitiveness.

Despite the high potential of data utilization for the manufacturing industry, expected by practitioners and scholars, manufacturing companies find it challenging to realize this potential. A 2016 survey on “Manufacturing Data Analytics” conducted by the Institute of Technology Management (ITEM-HSG) found that only about 11 percent of the collected data is actually used. Most companies lack the experience of data utilization and "do not know what to do with the data they have, let alone how to interpret them to improve their processes and products" (Kusiak, 2017, p. 23). Therefore, research is needed to support companies in identifying and implementing opportunities to better capitalize on their manufacturing data to improve quality, cost, and flexibility and thus, to achieve a competitive advantage.

Example: Data-based application (DBA) at an automotive supplier site

The company operates a company-wide LPS with five core principles: quality, delivery, cost, sustainability, and safety. A high level of overall equipment effectiveness (OEE) of bottleneck equipment is critical to the company to ensure deliverability as specified in service-level agreements with its customers. To minimize planned and unplanned downtime of critical production equipment, the company applies at least three different DBAs. First, the production equipment is continuously monitored by the application *Real-time Control* enabled through the installation of sensors. In case of any disturbances of the process, this application immediately informs the employee in charge, thus reducing the time until the problem is fixed. Second, based on sensor data, the DBA S*ystem Performance Measurement* calculates and visualizes the OEE over time. This allows the identification of trends as well as to benchmark a machine or process to a similar machine or process with the objective to identify opportunities for improvement. Third, based on historical and current machine condition data, the DBA *Predictive Maintenance* is used to derive a more effective and efficient maintenance plan. Thereby, unplanned machine breakdowns are minimized, and planned maintenance stops are reduced to a minimal level. Both effects lead to an increased OEE level.

Integrating new technologies is especially challenging for lean companies operating an LPS. Toyota, the founder of the Toyota Production System (TPS), which serves today as the blueprint for LPS, has tended to lag behind other companies in introducing new technology. New technology was not introduced before it had proven that it does not interfere with value-adding core processes, but provides a positive contribution to the existing system (Liker, 2004, p. 159).

The emergence of DBAs as part of SM and their potential to increase competitiveness raises three essential questions to companies operating an LPS. First, which DBAs exist in manufacturing and how can they support the companies' competitive priorities? Second, lean managers need to understand the impact of data utilization in manufacturing on the established lean practices to derive informed decisions on new investments into DBA. And third, implementing DBAs poses special requirements to the technical infrastructure but also to employees and to the organization. Understanding the key enablers to implement and use DBAs successfully is crucial to managers who are keen to achieve a competitive advantage for their companies by fostering the utilization of manufacturing data.

1.1.2 Scientific Relevance

The academic discipline of operations management (OM) is concerned with solving problems in the manufacturing environment (Peinado, Graeml, & Vianna, 2018, p. 374). Also, a central issue of scholars in the field of OM has always been to understand and describe the causes of competitive advantages of manufacturing companies (Barney, 1991; Hitt, Xu, & Carnes, 2016).

For example, the book *The Machine That Changed the World* by Womack et al. (1990) discusses the reasons for Japanese companies' superior quality and productivity and the resulting competitive advantage over U.S. competitors from 1960 on. This contribution is one of the most cited references in OM (Holweg, 2007, p. 420) and has made the concept of LM known to a broader audience. Since then lean thinking has made a significant impact both in industry and academia (Hines, Holweg, & Rich, 2004, p. 994), and is still an ongoing topic for future changes in the manufacturing environment (Prinz, Kreggenfeld, & Kuhlenkötter, 2018, p. 22). The fact that LM enjoys continuous popularity among managers of the manufacturing industry is enough justification for further research of the phenomena (Lewis, 2000, p. 976).

In recent years, the concept of SM has not only gained significant popularity in the industry, but also in academia[4] (Buer, Strandhagen, & Chan, 2018, p. 2924; Kusiak, 2018, p. 509). Tao, Qi, Liu, and Kusiak (2018, p. 1) link the concept of SM to data analytics and argue that data analytics offers tremendous opportunities for manufacturing companies. Lee, Lapira, Bagheri, and Kao (2013), Clegg and Powell (2013), and Kusiak (2017) consider data usage as a major factor of sustained manufacturing competitiveness. Thoben et al. (2017, p. 3) conclude that manufacturing intelligence is a compelling topic for practitioners and scholars worldwide. However, data integration into manufacturing also receives much attention; Tantik and Anderl (2016) express the need for further research in this area.

In the past, the two concepts of LM and SM have often been treated as separate subjects (Prinz et al., 2018, p. 21). Only recently has their relationship received increased attention in OM research (Rossini, Costa, Tortorella, & Portioli-Staudacher, 2019; Wagner, Herrmann, & Thiede, 2017). While the need for integrated research of SM and LM has been recognized, previous OM literature fails to address the integration of data utilization in LM.

This research seeks to address gaps in existing research by compiling a collection of DBAs in the manufacturing industry and by evaluating their potential to support lean practices.

---

[4] The increasing interest of the academic world in SM is reflected by the growing number of related scientific publications on the online database ScienceDirect. Number of search results per year for the term "smart manufacturing" on https://www.sciencedirect.com/ (23.06.18): 2000-2010: 20 articles; 2011-2015: 99 articles; 2017: 193 articles.

Also, it identifies key enablers for applying DBAs. Technical-oriented papers on data utilization from the SM stream mainly address technological challenges. However, as the history of LM and the computer integrated manufacturing (CIM) era (Kolberg, Knobloch, & Zühlke, 2016, p. 2853) shows, applying new tools and changing the way of working requires acceptance of employees and often involves organizational changes. Following the suggestion of Hirsch-Kreinsen et al. (2018, p. 181), this dissertation identifies organizational enablers as well as enablers regarding employees to complement technological enablers to implement and use DBAs successfully.

1.2 Research Gaps

The review of the relevant literature on LM, SM, and data utilization in manufacturing has revealed three research gaps.

First, from a practical perspective, many companies lack the experience to utilize manufacturing data. While most practitioners from the manufacturing industry believe that the use of data has a high potential, companies currently struggle to create value from the collected data. The academic literature documents a variety of use cases of DBAs; for instance, for predictive maintenance (Z. Li, Wang, & Wang, 2017), material flow management (Kolberg & Zühlke, 2015), and quality improvement (Gewohn, Usländer, Beyerer, & Sutschet, 2018). Also, scholars have contributed overviews on DBAs structured according to the underlying technology; for example, *big data application* (O'Donovan et al., 2015b), *data mining applications* (Harding et al., 2006), and *artificial intelligence applications* (Meziane, Vadera, Kobbacy, & Proudlove, 2000). Also, a variety of authors from the SM or industry 4.0 stream present potential use cases of DBAs but remain on a generic level. For instance, Lu (2017, p. 7) argues that "Industry 4.0 makes factories more intelligent, flexible, and dynamic by equipping manufacturing with sensors, actors, and autonomous systems," without specifying how sensors contribute to a more intelligent factory.

While the knowledge on DBAs exists and is documented in scientific literature, a comprehensive overview of DBAs relevant for the manufacturing industry is currently missing. This lack constitutes the first research gap. The author of this dissertation holds the view that a structured collection of manufacturing DBAs—including basic operating principles, objectives, and requirements—is valuable for scholars and practitioners. The latter will benefit from a scientifically-backed overview of DBAs to identify new opportunities to exploit manufacturing data.

While many publications on DBA use cases present the inherent technical challenges of the application, they almost universally neglect the aspect of employee acceptance of new tools and working procedures. However, considering the rich literature of LM (Friedli, Basu, Bellm, & Werani, 2013, p. 110; Liker, 2004, 36; Sanders, Elangeswaran, & Wulfsberg, 2016, p. 826), it is evident that employee involvement and qualification is a critical enabler for implementing new practices. This lesson learned from the lean journey should not be forgotten but kept in mind when introducing new technologies and tools. The lack of consideration of required non-technological enablers for implementing DBAs constitutes the second research gap.

Finally, although the literature of OM acknowledges the fundamental role of LM as a source of competitive advantage of manufacturing companies and characterizes data utilization as the fuel of the next industrial revolution (S. Yin & Kaynak, 2015, p. 144), research integrating both aspects is absent in current literature. This absence constitutes the third research gap.

1.3 Terms and Definitions

Providing clear definitions at the beginning of academic work is essential to ensure a shared understanding and to communicate ideas and findings accurately (Creswell, 2014, pp. 42–43). This section provides brief definitions of the following terms: Production/manufacturing, lean manufacturing, lean practices, smart manufacturing, and data-based application.

Production/manufacturing

Both terms are used interchangeably and refer to the combination and transformation of input factors such as material, utilities, and labor to finished products.

Lean manufacturing (LM)

Despite, or perhaps because of, the great attention lean production has received from academics as well as practitioners, the definition of lean production is still elusive (Pettersen, 2009, p. 127). In this dissertation, the following definition of LM by Shah and Ward (2007, p. 791) is applied: "Lean production is an integrated socio-technical system whose main objective is to eliminate waste by concurrently reducing or minimizing supplier, customer, and internal variability."

Lean practices

Womack and Jones (2003, pp. 15–26) have introduced five guiding principles to become a “lean company.” Lean practices are tools and methods that operationalize these five lean principles in order to achieve high customer value with minimal waste (Tortorella & Fettermann, 2018, p. 2). This dissertation refers to a collection of lean practices, resulting from a comprehensive literature review on LM by Shah and Ward (2003).

For the reason of clarity, lean practices are capitalized in this dissertation, (e.g., Preventive Maintenance*).* The introduction and use of abbreviations was deliberately dispensed with due to the number of lean practices.

Smart manufacturing (SM)

Based on a literature review on SM, Thoben et al. (2017, p. 6) conclude that a variety of different definitions for SM exist. The definition used in this dissertation is provided by O'Donovan et al. (2015b, p. 2): "Smart manufacturing can be considered the pursuit of data-driven manufacturing, where real-time data from sensors in the factory can be analyzed to inform decision-making. More generally, smart manufacturing can be considered a specialization of big data, whereby big data technologies and methods are extended to meet the needs of manufacturing."

The term *smart manufacturing* is used synonymously with the term *industrie 4.0* or *industry 4.0* as both refer to the same concept (Wuest, Weimer, Irgens, & Thoben, 2016, p. 23).

Data-based application (DBA)

The term *data-based application* is applied in this work as an umbrella term for several distinct applications from the areas of data mining (DM), machine learning[5] (ML), mathematical optimization, and simulation. A DBA is a tool or system that is utilizing data to reach a particular purpose. The mere data collection without further utilization is not a DBA as it serves no purpose. A DBA is located at the end of the manufacturing data lifecycle consisting of data collection, data processing, and data application. Examples for DBAs are track and trace, performance measurement of the production system, and data-based predictive maintenance. Chapter 4.1.1 provides a more detailed definition by presenting four criteria for DBAs.

As shown later, some DBAs include data analytics while others do not. Analytics is defined in this dissertation as "a scientific process of logical-mathematical transformation of data to improve decision-making" (Blum & Schuh, 2017, p. 258). The term *data analytics DBA* refers to those sophisticated DBAs that rely on advanced data analytics.

For the reason of clarity, DBA categories and DBAs are written in italics and are capitalized in this dissertation, (e.g., *Predictive Maintenance).* The introduction and use of abbreviations was deliberately dispensed with due to the number of DBAs.

[5] The term ML is used synonymously with the term AI

1.4 Research Scope and Objectives

This chapter outlines the research scope, the research questions, and the objectives.

1.4.1 Research Scope

Data utilization is of increasing importance in many industries, especially large internet companies such as Alphabet (the mother company of Google), Amazon, and Facebook that are well-known for their data-driven business models. Provost and Fawcett (2013b) present several use cases of data utilization. For instance, in retail for predicting customer demand or for communication companies to estimate the likelihood of customer churns. A full discussion of data utilization in general lies beyond the scope of this work. The scope is limited from two perspectives. First, only DBAs relevant to the manufacturing industry are considered. Second, although there are use cases of data utilization regarding customer and supplier integration, the scope of this work is set to the production facility.

Regarding the relationship between LM and SM, the literature summarizes three different perspectives. While some scholars (Bick, 2014; Künzel, 2016; Metternich, Müller, Meudt, & Schaede, 2017; B. Wang et al., 2016) advocate the view that LM enables and facilitates SM, others (Kolberg & Zühlke, 2015; Wagner et al., 2017) take the perspective that SM advances LM. The third group of scholars (Sanders et al., 2016; Tortorella & Fettermann, 2018) reports a mutually beneficial relationship between both production paradigms, without stating the direction of support. The research at hand takes the second perspective by researching the question of how DBAs, which are closely linked to SM, may advance lean practices and thus, support LM.

1.4.2 Research Questions and Research Objectives

The research questions and research objectives are derived from the deficiencies of current research as identified in chapter 1.2. Each identified research gap is addressed by a sub-research-question (SRQ). The combined findings of the three SRQs allow to provide a comprehensive answer to the main-research-question (MRQ) stated in Table 1.

Table 1: Research questions

MRQ	How can manufacturing companies be enabled to implement data-based applications to support lean practices?
SRQ 1	Which data-based applications exist in manufacturing and what are their objectives?
SRQ 2	What are key enablers to apply data-based applications?
SRQ 3	How can data-based applications support lean practices?

The MRQ intends to investigate how companies operating an LPS can benefit from the potential of data utilization to support the effectiveness of established lean practices. More effective implementation of lean practices supports the key objectives of every LPS: high quality, high deliverability, and low costs, and thus contributes to the company's competitiveness.

To answer the MRQ, three SRQs with distinct foci have been formulated. SRQ 1 provides the foundation for further research by consolidating existing use cases of DBAs in manufacturing in a structured overview. Furthermore, SRQ 1 includes a description of the underlying operating principle of each DBA, the objectives, and specific requirements for their implementation.

SRQ 2 builds on SRQ 1 by consolidating and generalizing the specific requirements. Being aware of the central importance of human and organizational factors for implementing new tools and work methods, SRQ2 broadens the focus of enablers and also takes necessary organizational enablers and employee enablers into consideration.

Finally, SRQ 3 links the concept of data utilization in manufacturing to LM. Referring to research on the interaction of SM and LM, Mayr et al. (2018, p. 623) criticize that most current publications address the issue on a general level, while the link to a particular lean practice is often missing. The dissertation at hand considers this remark by selecting the level of lean practices as most appropriate to evaluate the impact of DBAs on LM. More specifically, SRQ3 systematically evaluates the impact of DBAs on established lean practices by a pairwise evaluation. A key question of the evaluation is whether DBA will allow lean practices to be more effectively applied, or whether DBA might serve as a substitute, thus decreasing the importance of a particular lean practice.

1.5 Research Design

The following chapter outlines the research design of this dissertation. Chapter 1.5.1 presents the conceptual background and discusses the basic philosophy underlying the research. Chapter 1.5.2 presents the research process, and chapter 1.5.3 the applied research methodology. Chapter 1.5.4 depicts the research framework, and chapter 1.5.5 closes by introducing the research theory.

1.5.1 Conceptual Background

The research follows the understanding of business studies as applied science as introduced by Ulrich (1984). According to this understanding, applied science’s intent is to develop rules, models, and methods for practical action based on findings of theoretical and basic research. In contrast to natural science, which observes and explains an existing reality, social science seeks to develop and create a new reality. Furthermore, social science is distinguished from natural science by its interdisciplinary character. While natural science usually has a very distinct focus, social science needs to integrate the social context (Ulrich, 1982, p. 5). Social science recognizes the complexity of social systems and abandons the requirement of total control (Ulrich, 1984).

Following Ulrich (1982, pp. 3–4), the problems researched by applied science originate from practice and therefore are defined outside the academic world. Consequently, the value of research of applied science is not defined by the validity of a theory but rather by the usefulness of models and rules for practical applications. In accordance with Ulrich's understanding of applied social science, this research is motivated by a practical problem with the objective to provide scientifically backed suggestions to the practical world on how to use DBAs to support LP.

Creswell (2014, p. 5) differentiates four philosophical worldviews in research: postpositivism, constructivism, transformative, and pragmatism. The underlying philosophical worldview of this research is in line with Ulrich's perception of management studies. The research objective is real-world practice-oriented and problem-centered. Following the classification of Creswell (2014, p. 6), the author of this research would position himself as a pragmatist. In contrast to representatives of the postpositivism and constructivism worldview, the pragmatic worldview is less constrained in selecting appropriate research methods and often uses both quantitative and qualitative data (Creswell, 2014, pp. 10–11).

1.5.2 Research Process

To ensure a structured and heuristic approach, this research follows the iterative learning process as described by Kubicek (1977), Tomczak (1992), and Gassmann (1999). The generic process is shown in Figure 1.

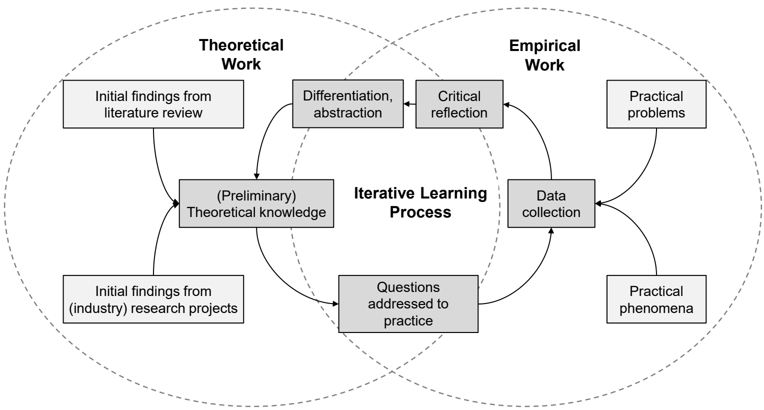

Figure 1: Iterative Learning Process (adapted from Gassmann (1999, p. 13) and Tomczak (1992, p. 84))

Grounded on preliminary theoretical knowledge, questions are derived for research in practice. Preliminary theoretical knowledge is gained through literature review and initial discussions with research partners to get a firm understanding of the problem and to formulate meaningful questions to reality. Data collection may include the compilation of qualitative (e.g., interviews and case studies) as well as quantitative data (e.g., survey, access to historical data).

Based on collected data from the practice, a new picture of reality is drawn. The new knowledge is critically reflected and enriches the existing theoretical knowledge of the researcher, resulting in a differentiation, abstraction, or change of perspective of the preliminary understanding of the research problem. The iterative learning process supports the researcher to reduce the risk of incorrect assumptions by reevaluating the current state of knowledge iteratively (Kubicek, 1977; Tomczak, 1992). Moreover, Kubicek (1977) and Tomczak (1992) argue that the iterative learning process is well suited for research addressing topics with limited existing knowledge.

In the context of this research, preliminary knowledge led to the assumption that emerging advanced technologies will have an impact on the manufacturing industry and therefore also on LPSs. This assumption was integrated into a study on the status quo and potential directions of development of LM, called “Lean2020 – The Future of Operational Excellence” (Macuvele et al., 2018). To get an idea of the impact of new technologies on LM, we integrated aspects such as current and planned technology utilization as well as questions on data collection and utilization. The study results indicate that on average the participating companies see the highest potential of (big) data collection and analytics to support LM, compared to other trends of digitalization. Furthermore, the study revealed that successful companies invest considerably more into data collection and data analytics (see chapter 3).

Critical reflection of the collected data led to a more detailed scope of the research. Coming from the very broad scope of the interaction of SM technologies and LM in general, the subsequent research focused on the impact of DBAs on LM practices. Following this specialization of the research scope, a comprehensive literature research was conducted with three objectives: (1) understand how SM creates a technological basis for DBAs, (2) identify and cluster existing use cases of DBAs within the manufacturing industry, and (3) identify examples where LM benefits from the integration of SM. Based on the findings of the literature review, conceptual research was conducted to derive propositions of DBA support for lean practices.

In the third iteration of data collection, case study research in collaboration with three companies and two expert interviews with senior academics was conducted to complement the findings of the quantitative survey and the literature review.

1.5.3 Research Methodology

The type of research question needs to be consistent with the selected research methodology. As this research follows an inductive approach, that is to conclude from a limited number of observations to general themes (Creswell, 2014, p. 4; Tomczak, 1992, p. 77), this research follows a qualitative approach, including case study research and expert interviews.

Based on the review of case study literature, Gerring (2009, p. 37) defines case study research as "an intensive study of a single unit or a small number of units (the cases), for the purpose of understanding a larger class of similar units (a population of cases)."

Case study research is used as a primary mean of data collection for three reasons. First, case study research is qualified to assess problems originating from the practical world because the research of real-world problems of industry results in a higher practical relevance of the findings. Managers value case studies as they are interested in learning from experience. This includes, for instance, arising problems, challenges, and approaches—not only from their own company, but also from other companies—thus enabling inter-organizational learning (Gassmann, 1999, p. 11).

Second, Voss, Tsikriktsis, and Frohlich (2002, p. 198) argue that case study research is especially useful in research areas with little existing knowledge and unclear definitions of key constructs. This applies to the research at hand, as companies and scholars are just starting to explore the opportunities of DBAs in manufacturing.

Third, case studies are beneficial for researching rather explanatory research questions beginning with *how* and *why* (R. Yin (2009, p. 36)*)* or *what* (Creswell, 2014, p. 140). As the MRQ and two of three SRQs are *how* and *what* questions, qualitative case study research is considered appropriate for this research.

Voss et al. (2002, p. 195) highlight two additional advantages of case study research. First, it is not constrained as questionnaire-based surveys and allows questions to be answered in more depth, leading to a more comprehensive understanding of the nature and complexity of the unit of analysis. Second, the researched item can be investigated in its natural setting, thus increasing the relevance of the observations. The reasons for applying case study research outlined above contributed to the fact that case research is described by Voss et al. (2002, p. 195) as the "most powerful research method in operations management."

In addition to case study research, semi-structured expert interviews, which are also associated with qualitative research (Trinczek, 2009, p. 204), are part of the data gathering process. Expert interviews are especially attractive in situations when the expert can act "as a surrogate for a wider circle of players" (Bogner, Littig, & Menz, 2009, pp. 1–2) and therefore increase the effectiveness of data collection.

1.5.4 Research Framework

In general, academic problems are multilayered and multifaceted. Often there is a complex mesh of elements and relationships. To structure the research, the use of a frame of reference is advised (Wolf, 2011, p. 37). A frame of reference, or research framework, is first a descriptive model that provides orientation and structures the research process. Also, a research framework indicates relationships between the elements but does not explain the type or cause of the relationship. By presenting the core elements and the essential relationships of the unit of analysis, a research framework serves two purposes. First, it forces the scholar to clarify his or her research scope and objectives. Second, it facilitates the communication and discussion of research scope and objective with other academics as well as project partners. (Wolf, 2011).

The research framework of this dissertation is depicted in Figure 2.

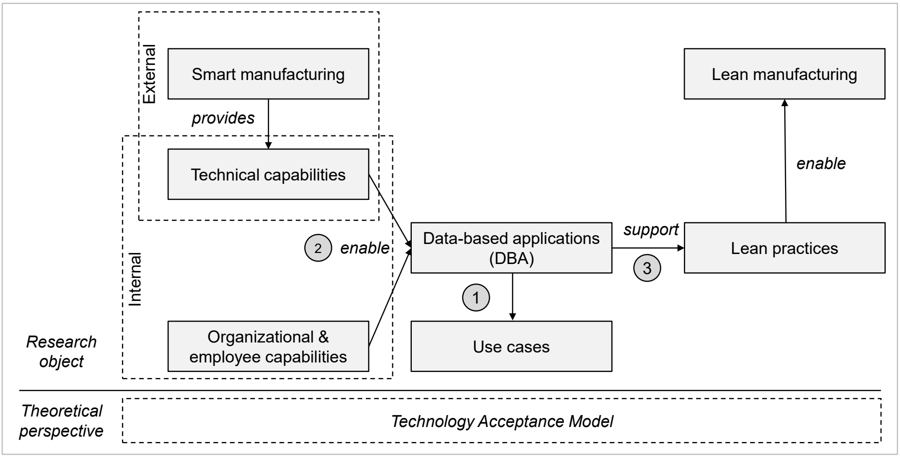

Figure 2: Research framework

Academic literature has treated the field of SM and LM, in general, separately. The research framework indicates that according to the authors' initial theoretical knowledge, DBAs may serve as a *transmission belt* that transforms the technological advances of SM to increased effectiveness of lean practices of LM, and therefore links both production paradigms. The framework visualizes the assumption that DBAs have a positive impact on certain lean practices. Finally, since lean practices are enablers of LM, more effective implementation of lean practices supports LM to achieve its objectives. All three SRQ are positioned within the research framework: (1) indicates the collection of DBA use cases in manufacturing, (2) indicates the identification of main enablers, and (3) indicates the systematic evaluation of the support potential of DBAs on lean practices.

1.5.5 Research Theory

Kerlinger & Lee (2000, p. 11, cited in Levy & Ellis, 2006) define theory as “a set of interrelated constructs (concepts), definitions, and propositions that present a systematic view of phenomena by specifying relationships among variables, with the purpose of explaining and predicting the phenomena.” Depending on the research methodology, the research theory may serve a different function. In quantitative research, the theory proposes an explanation for the relationship between the tested variables. In qualitative research, a theory can either be a product of the research or the research theory serves as a lens for the researcher to look at the problem (Creswell, 2014, p. 51).

Technology Adoption Theory

Acceptance of employees is a fundamental requirement for every successful introduction of new technologies and new working methods (Kolberg et al., 2016, p. 9). Acceptance is this context is defined as “an antagonism to the term *refusal* and means the positive decision to use an innovation” (Taherdoost, 2018, p. 961). Due to the importance of users’ acceptance for technology, scholars search for fostering or hampering factors (Taherdoost, 2018, p. 960). Table 2 provides an overview of common technology adoption theories and models grouped by their purpose

Table 2: Common technology adoption theories and models (adapted from Hillmer, 2009, p. 18)

Theory	Diffusion Theories	User Acceptance Theories	Decision-Making Theories	Organization Structure Theories
Focus	Technology, the environment, and the using organization	Rational employee interest	Rational Choice Theory/ Game Theory	Strategic organizational interest
Example	Diffusion of Innovation Theory DOI (Rogers, 1983)	Theory of Reasoned Action TRA (Ajzen & Fishbein, 1973)	Decision-Making under Uncertainty	Creative Destruction Theory (Schumpeter 1942)[6]
	Technology Lifecycle Theory (Moore, 1995)	**Technology Acceptance Model** (Davis, 1989)	Technology Lifecycle Theory	Disruptive Technology Theory (Hillmer, 2009)
		User Acceptance of Information Technology (Venkatesh, Morris, & Davis, 2003)	Change Management	

[6] Cited in Baaij, Greeven, and van Dalen (2004)

This dissertation intends to identify enabling factors that allow companies to implement and use DBAs successfully. Ensuring employees' acceptance of DBAs is crucial for their successful use. Employees might perceive DBAs as complex, intransparent, and as a potential driver of job reductions on the shop floor, thus threatening the acceptance of this new approach. The *technology acceptance model* is used to find conditions under which employees are more likely to accept and use new technologies. Hence, it is an appropriate theoretical lens to guide this research, especially regarding the intention to identify challenges and enablers to implement DBAs.

Technology Acceptance Model (TAM)

The TAM goes back to Davis (1989) and is probably the most often applied theory within the group of user acceptance theories. It is based on the *theory of reasoned action* but has been adapted to fit the context of information technology. The theory of reasoned action originates in the field of social science and argues that behavior is a function of behavioral intentions (Hillmer, 2009, pp. 17–19). The TAM postulates that the actual use of new technology is driven by a person's *behavioral intentions to use (BI)* the technology. The BI, in turn, is influenced by two aspects: first, by the *attitude towards using (A)* the technology and second, directly by the *perceived usefulness (U)* of the technology. Attitude towards using technology is influenced on the one hand by the perceived usefulness and the *perceived ease of use (EOU)*, while both of these factors are influenced by external variables which are not further specified (see Figure 3). Davis (1989, p. 320) defines *perceived usefulness* as “the degree to which a person believes that using a particular system would enhance his or her job performance” and *perceived ease of use* as “the degree to which a person believes that using a particular system would be free of effort.”

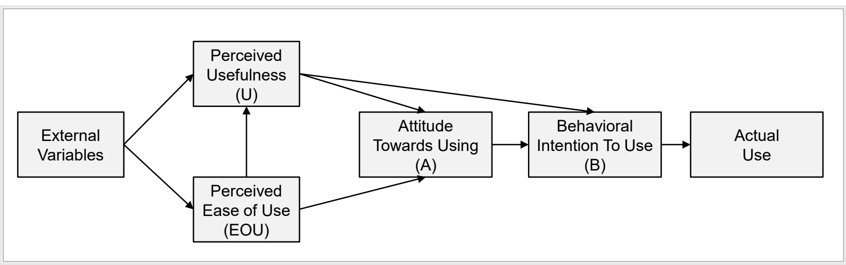

Figure 3: Technology Acceptance Model(TAM) (based on Davis, 1989)

Employee acceptance is essential for introducing new technologies and working methods successfully. Therefore, employee acceptance is an important aspect to address when searching for key challenges and enablers for the implementation of DBAs. The TAM explains the actual use of technology by presenting factors that positively influence the behavioral intention to use a technology. Similarly, the qualitative studies in chapter 5 seek to find positive and negative influencing factors to employees' acceptance of new DBAs. However, to avoid biasing the answers of interview partners, the key influencing factors of the TAM, *perceived usefulness (U)* and *perceived ease of use (EOU)*, are not presented during the interview.

Nevertheless, the model is very useful to serve as a reference to the findings of drivers of employee acceptance of the qualitative studies. Therefore, chapter 6.3.2 contrasts the key factors of employee acceptance of the TAM to the factors identified in chapter 5. By comparing the findings, the explanatory power of the TAM is critically evaluated and, if necessary, propositions for extensions made.

1.6 Thesis Outline

This dissertation has seven chapters and comprises quantitative, conceptual, qualitative, and concluding parts.

Chapter 1: Introduction

Chapter 1 discusses the practical and scientific relevance, and identifies three gaps in existing research. Based on these gaps, one MRQ and three SRQ questions are derived. Also, this chapter introduces the research design to include the conceptual background, research process, research methodology and framework, and the research theory.

Chapter 2: State of Research

Chapter 2 introduces the literature review process and provides an overview of the state of the art of two research streams. It introduces the concept of LM, followed by an overview of the emerging paradigm of SM, and a discussion of current research on the interplay of LM and SM. Furthermore, state-of-the-art techniques of data analytics and the role of data in SM are briefly discussed.

Chapter 3: Smart Manufacturing Technologies in Lean Manufacturing – Quantitative Study Results

This chapter presents selected results from a quantitative survey on the future of LM conducted by the author in 2017 (Macuvele et al., 2018). On one hand, general findings are shown, and on the other hand, patterns of successful practice companies are identified and discussed. The chapter closes with implications for the following chapters.

Chapter 4: Data-based Applications in Lean Manufacturing

Chapter 4 consists of two parts. The first part addresses SRQ 1 by conducting a comprehensive literature review to identify use cases of DBAs in the manufacturing industry. The findings are then consolidated in a structured overview. It also summarizes key requirements to implement DBAs, which can be derived from the literature review. Thus, the first part also addresses SRQ 2.

The second part addresses SRQ 3 and comprises a conceptual approach to derive a systematic evaluation of the potential impact of DBAs on lean practices. The DBA—Lean Practice Impact Matrix—is presented and identified support potentials discussed.

Chapter 5: Qualitative Studies

The literature review and conceptual part of chapter 4 is complemented with qualitative study results in chapter 5. This chapter comprises the methodology and the results of the applied case study research as well as the results of expert interviews. The qualitative studies focus on organizational enablers and employee enablers to implement DBAs and thus complement the findings of chapter 4 to answer SRQ 2.

Chapter 6: Consolidation of Findings

Chapter 6 consolidates the findings from chapter 3, chapter 4, and chapter 5 to answer the three sub-research-questions. It comprises three parts. The first part summarizes the findings of DBAs in manufacturing in general, including an overview of DBA and their objectives (SRQ 1) as well as the key challenges and enablers to apply DBAs (SRQ 2).

The second part takes the perspective of a company operating an LPS. It presents the main benefits of data utilization for LM, the potential support of DBAs to lean practices (SRQ 3), as well as potential threats that are specific to lean companies.

The third part formulates three theoretical implications by abstracting the findings from the specific use cases. First, it introduces a DBA value model that distinguishes three value levels of DBAs. Second, it presents the *ROI Dilemma of DBAs*, which builds on the previous research results and might explain the hesitant approach of many manufacturing companies toward DBAs. Finally, the research theory TAM is critically evaluated and propositions for extensions are made.

Chapter 7: Conclusion and Outlook

The final chapter answers the MRQ by summarizing the results of the research. Also, it discusses the theoretical and practical contributions of the research and presents some limitations arising from the qualitative studies. Finally, an outlook is given on the potential for further research.

2 State of Research

This chapter summarizes the state of research for the literature streams of LM and SM. Furthermore, recent work on the integration of both production paradigms is outlined. To understand the foundation of many of the DBAs introduced in chapter 3, the last section of this chapter introduces basic methods of state-of-the-art data analytics.

2.1 Literature Review Process

A review of the current state of the research in the academic literature is of utmost importance for every scientific work. Levy and Ellis (2006, p. 183) and Randolph and Justus (2009, p. 2) present several reasons for conducting a literature review prior to the research. First, researchers need to understand the state of the art of current knowledge to answer the question of what is already known and to identify gaps where more research is needed. Second, existing theories provide a sound theoretical foundation for the research. Third, by establishing the state of existing research, the advances through new research can be documented, and the proposed research justified.

The literature review of this dissertation follows the systematic approach presented by Levy and Ellis (2006) comprising three phases: literature input, literature processing, and literature output.

Literature Input

The first phase includes all actions to define and search the relevant literature of selected research streams. In the context of this research, the two research streams of LM and SM have been analyzed. Besides considering the streams separately, current research on the interplay between both production paradigms is presented. Finally, state-of-the-art techniques of data analytics are briefly introduced.

The definition of relevant input literature follows the taxonomy of literature reviews of Cooper (1988, p. 109). According to Cooper, the following six characteristics are used to classify literature reviews (see Table 3): *focus (1)*, *goal (2)*, *perspective (3)*, *coverage (4)*, *organization (5),* and *audience (6).*

The first characteristic is the focus (1) of the reviewer. “The focus of a review concerns the material that is of central interest to the reviewer” (Cooper, 1988, p. 108). This review focuses on research outcomes, research methods, and in particular on practices and applications. The goal (2) describes what the author intends to achieve with the review. The objective of this review is to integrate and synthesize existing contributions and to identify central issues of the research field. Regarding perspective (3), the author takes a neutral representation. Given the abundance of academic publications on LM in the last 30 years, the review intends to provide a representative coverage (4) rather than an exhaustive one on LM. Consequently, standard references and review articles are integrated rather than all current publications on LM.

Although the field of SM is much younger, due to its interdisciplinary character between information system research and OM research, a plethora of research articles are available too. Therefore, the literature review targets an exhaustive coverage of the field of SM but relies on selective citation of current contributions. The same applies to the interplay of LM and SM. Concerning current data analytics techniques, only central concepts are covered.

The organization (5) of the review is conceptual; that is, bringing together works relating to the same abstract concept. It is also methodological, as it conflates works applying similar methods. The target audience (6) consists of general scholars and especially practitioners from the manufacturing industry.

Table 3 shows the taxonomy of literature reviews and highlights the characteristics of this review in *italics*.

Table 3: Taxonomy of literature reviews (Cooper, 1988)

Characteristic	**Categories**			
1. Focus	*Research Outcomes*	*Research Methods*	Theories	*Practices or Application*
2. Goal	*Integration*	Criticism	*Identification of Central Issues*	
3. Perspective	*Neutral Representation*	Espousal of Position		
4. Coverage	Exhaustive	*Exhaustive with Selective Citation*	*Representative*	*Central or Pivotal*
5. Organizational	Historical	*Conceptual*	*Methodological*	
6. Audience	Specialized Scholars	*General Scholars*	*Practitioners* or Policy Makers	General Public

The literature search includes high-quality journal papers and, if necessary, peer-reviewed conference contributions. The quality of journals was, so far applicable, assessed by consulting the VHB ranking of the categories *production management* and *technology, innovation* and *entrepreneurship*[7] as well as the ranking on IS journals and conferences provided by Levy and Ellis (2006, pp. 186–187).

Although, in general, the quality and academic rigor of conference proceedings is perceived lower compared to journals (Levy & Ellis, 2006, p. 187). Peer-reviewed conference articles have also been included for two reasons. First, conference papers tend to be more recent (vom Brocke, Simons, Riemer, Niehaves, & Platfaut, 2015, p. 210), thus reflecting the most current developments. Second, conference contributions often follow a more practice-oriented approach in an industry setting, therefore the results are often less abstract and high-level but can be used to draw a realistic picture of state-of-the-art application of a technology in the industry.

Journals and conference contributions are obtained from the databases *EmeraldInsight, EBSCOhost, Web of Science, ScienceDirect, and ProQuest.* A list of journals identified in the review and considered for this research is presented in Appendix A.

At first, publications were identified by keyword search. Following vom Brocke et al. (2015) the search parameter is a combination of search terms (e.g., “lean”), search operators (e.g., “AND”), and search fields (e.g., abstract, title). Torraco (2005, p. 360) points out that “learning about the literature and how it was obtained, including the keywords and databases used, is of particular interest to readers, who may wonder if the literature they are familiar with was examined” and therefore asks for documentation of the search parameters.

Table 4[8] shows the search terms and search operators applied for the literature review of the research areas LM, SM, and data analytics. The literature on the interplay of LM and SM is covered as part of the general SM literature.

[7] German: Produktionswirtschaft and Technologie, Innovation und Entrepreneurship

[8] A different set of search terms is used and presented in chapter 4 to identify use cases of DBAs.

Table 4: Applied search terms and search operators for literature review

Research area	Search term 1	Operator 1	Search term 2	Operator 2	Search term 3
LM	Lean TPS Toyota Production System	AND	production manufactur* technology	AND	Practice* ""
SM	Smart Digital	AND	factory manufactur* technology	AND	Data ""
	Industrie 4.0 Industry 4.0	AND			
Data analytics	Data	AND	analytics analysis mining visualization optimization	AND	factory manufactur* production techniques methods

As an example to read: The first possible combination of search terms and operators covers literature that includes the terms lean production or lean manufactur*ing*, or lean manufactur*er* or lean technology in combination with the term practic*e* or practic*es* or without practices (indicated by “”). After eliminating duplications and initial screening for relevance based on the abstract, the initial keyword search resulted in 111 publications. For more information on search parameters and results, see Appendix B.

The search parameters were used as the search criteria in the five online databases presented before. It has to be noted that the input masks of the literature databases are not fully standardized, hence the search parameters have been adapted slightly to fit the syntax of the database. The general search criteria are as follows: only peer-reviewed papers with full-text access and German or English language have been considered. The search terms have been searched within *abstract and title*, or within *everything but full text*, depending on the database. The search period covered 11 years from 01.01.2008 until 31.12.2018.

Literature Processing

According to Levy and Ellis (2006), literature processing comprises six steps: first, knowing the literature. The scholar needs to demonstrate that he or she has read the literature and is able to extract relevant information from it. This includes activities such as listing, describing, and defining. Second, comprehend the literature. In addition to repeating information from articles, this step requires the ability to differentiate, interpret, contrast, and summarize the information. Third, apply the literature. This step mainly includes activities such as classifying and relating. A common approach is first to identify a general concept and then assign the respective literature to the right category. Fourth, analyze the literature. This step needs to answer why a piece of information taken from the literature is relevant. Fifth, synthesize the literature. Synthesizing is combining the gained knowledge to an integrated whole. Sixth, evaluate the literature. The final step comprises activities such as assessing information, concluding, and explaining results.

The purpose of the processing order is to support researchers in transforming a large number of unrelated publications to an effective literature review. It serves as guidance for the following literature review. Besides building on the content of literature identified in the literature input phase, the literature serves as a basis for backward and forward search. Backward search refers to reviewing the references of already identified literature to identify articles relevant to the research area that have been missed during keyword search (vom Brocke et al., 2015, p. 216). Forward search is the opposite approach, which is including literature that references the same articles already identified as relevant (Webster & Watson, 2002, p. 16). Backward search has proven to be valuable to broaden the literature base. In total, 28 publications were added to the literature base through backward search, while only four were added through forward search.

Literature Output

Based on the searched and processed literature, the final step comprises the actual writing of the literature review. The results of the literature review are used to outline the current state of research in the following sections and to identify the use cases of DBAs in chapter 4.

2.2 Lean Manufacturing

LM is characterized by scholars as the "most influential manufacturing paradigm of recent times" (Holweg, 2007, p. 420). A recent survey confirms its relevance not only for today's but also for the future world of manufacturing (see chapter 3). This chapter introduces LM with a special focus on widely established lean practices.

Chapter 2.2.1 presents the origin of LM and seeks to provide a definition, acknowledging that a universal definition of LM does not exist. Chapter 2.2.2 describes the fundamental components of the TPS which serves as a blueprint for current LPSs. In chapter 2.2.3, a selection of 10 widely established lean practices is presented and the lean practices described. Chapter 2.2.4 finally discusses the role of technology in LM and explains the seemingly hesitant approach of companies to integrate new technology into their LPS.

2.2.1 Origin and Definition

2.2.1.1 Origin

The term *lean* was introduced the first time by Krafcik (1988, p. 41). The term was coined during an extensive five-year comparative study of more than 50 production systems worldwide. Krafcik introduced the term lean[9] in contrast to buffered. Buffered referred to the system of western mass production at that time with high levels of safety inventory. Womack et al. (1990) adopted the term in its world-known publication *The Machine That Changed the World* and thus made the term lean known to a broad audience of scholars and practitioners.

The concept of lean originated from the TPS. After World War II, the Japanese automotive market was very small and capital extremely scarce (Holweg, 2007, p. 421), therefore copying the western model of capital intensive large batch production was not an option for Toyota. As a result, Toyota developed a very different approach to manufacturing, known today as TPS or LM.

2.2.1.2 Definition

LM has received great attention from scholars as well as practitioners. Despite, or perhaps because of this fact, the definition of LM is still elusive. Shah and Ward (2007, p. 786) agree and point out that discussions with managers, consultants, and academics quickly demonstrate the absence of a common definition of LM. One reason for the absence of an accepted definition is, according to Hines et al. (2004, p. 1005), the fact that lean has evolved and is still evolving. Therefore, any definition

[9] Originally, the term fragile was used, but Krafcik felt that fragile had a negative connotation and decided to use the term lean instead (Holweg 2007, p. 426).

would only be valid for a certain point in time. Although no unified definition exists, some commonalities between a variety of definitions can be found.

Table 5 depicts seven commonalities of selected contributions on LM.

Table 5: Commonalities of LM definitions

Characteristic and elements of LM	**Key references**									
	1	2	3	4	5	6	7	8	9	10
1. The primary objective is waste reduction	●	●	●	●	●	●	●	●	●	●
2. Customer orientation	●	●	●				●	●	●	●
3. Principle orientation	●		●	●		●	●		●	●
4. Implementation through lean practices	●	●	●		●	●	●	●	●	●
5. Human focus	●		●	●	●	●				●
6. Continuous improvement (CI)	●	●	●	●	●	●	●	●	●	●
7. System approach	●	●	●	●	●	●	●			●

References: 1. Liker (2004), 2. Shah and Ward (2003), 3. Womack and Jones (2003), 4. Lander and Liker (2007), 5. Friedli and Bellm (2013), 6. Toyota Material Handling (n.d.), 7. Bertagnolli (2018), 8. Duarte and Cruz-Machado (2013), 9. Pettersen (2009), 10. Tortorella and Fettermann (2018)

First, the primary goal of LM is to *eliminate all kinds of waste* during the value creation process. The concept of waste reduction is linked to the second commonality, *customer orientation*. Value has to be defined from a customer's perspective. Therefore, all activities that are not contributing to generate value from the customer's perspective are potentially wasteful. The third commonly accepted lean characteristic is *principle orientation*. As Womack and Jones (2003) note, LM follows the five principles of *(1) define value from the customer perspective*, (2) *identify the value stream*, (3) *flow*, (4) *pull*, and (5) *strive for perfection*. Fourth, the principles of LM are realized with the help of *lean practices*, such as preventive maintenance. Fifth, the TPS and its descendent LM follows a human-centered approach and considers the employees as the most important asset instead of a cost factor. Employees are trained regularly to keep a high level of qualification and are encouraged to contribute to the CI of the production system. Sixth, all reviewed scholars regard CI as a key element of LM. Finally, most scholars highlight the fact that the TPS or LM is a holistic and multi-dimensional system comprising technical as well as cultural aspects.

This dissertation does not claim to provide a universal definition of lean. However, the seven commonalities described above are considered good indicators of the essence of LM. As a working definition, the following definition of Shah and Ward (2007, p. 791) is used in this work:

“Lean production is an integrated socio-technical system whose main objective is to eliminate waste by concurrently reducing or minimizing supplier, customer, and internal variability.”

2.2.2 The Toyota Production System

“All we are doing is looking at the time line from the moment the customer gives us an order to the point when we collect the cash. And we are reducing that time line by removing the non-value-added wastes.”

(Ohno, 1988, ix)

Toyota suffered less from the economic downturn in the aftermath of the first oil crisis in the early 1970s and recovered faster than its western mass production-oriented competitors (Lander & Liker, 2007, p. 3681). In the following period, Toyota increasingly gained market share and was considered the world's most efficient car manufacturer (Womack et al., 1990, p. 49). According to several scholars, the principal cause for Toyota's competitive advantage was the TPS (Abernathy & Clark, 1981; Holweg, 2007; Lander & Liker, 2007; Shimokawa, 2010; Wheelwright & Hayes, 1985). Although playing a significant role in Toyota's success, the TPS was not formally documented before 1965, when the Kanban system was introduced to its suppliers (Holweg, 2007, p. 423).

Following an official publication of Toyota MH[10] (n.d.), the TPS is a production system that continuously strives to eliminate waste in natural, human, and corporate resources. Therefore, the TPS empowers employees to improve processes and optimize quality permanently.

The TPS comprises a set of values, knowledge, and procedures, which are shown in the TPS House in Figure 4.

The roof of the TPS house consists of its goals, which are high quality, low costs, short lead times, and the highest level of safety. All goals are linked to the main objective of eliminating waste. The fundament of the house is formed by the two elements *Heijunka* and *Standardization.* Heijunka means leveling out the schedule to avoid

[10] Toyota Material Handling

unevenness (Mura) and aims to reduce variations of production volume and mix (Liker, 2004, p. 113). Standardization is necessary to make tasks repeatable but is also the foundation for CI. Only after having a process standardized and stabilized can it be improved (Liker, 2004, p. 142).

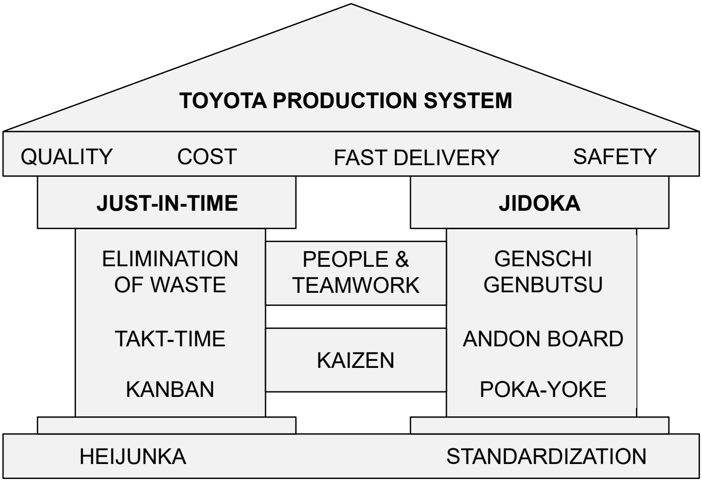

Figure 4: Toyota Production System (adapted from Liker, 2004, p. 33 and Toyota MH, n.d., p. 5)

The first pillar of the house is *just-in-time* (JIT) production. It comprises a set of principles and tools to deliver the right parts in the right quantity at the right time, and allows Toyota to respond quickly to shifts in customer demand (Shah & Ward, 2007, p. 788; Toyota MH, n.d., p. 8). The second pillar of the TPS house is *Jidoka*. Jidoka is often translated to *automation with a human touch* (Liker, 2004, p. 16) and means to ensure quality right at the moment of production instead of subsequent quality control. In the TPS, quality and problem solving is everybody's responsibility (Towill, 2007, p. 3620; Womack et al., 1990, p. 54).

Finally, Liker (2004, p. 33) puts *People and Teamwork* and *CI* at the center of his visualization of the TPS house. Toyota encourages its workers to embrace change and continuously improve their working space, even at the risk of making their own workplace obsolete. In return, Toyota offers a high level of employment security and invests in training and skill development of its workers. CI, or *Kaizen*, is the principle of permanently challenging the status quo to improve it. Kaizen requires employees to understand and think about processes (Toyota MH, n.d., p. 12).

The TPS house visualizes a set of principles and tools, but the TPS is more than just a collection of principles and tools. To quote the former CEO of Toyota, Fujio Cho (Liker & Morgan, 2006, p. 9): “The key to the Toyota Way and what makes Toyota stand out is not any of the individual elements. . . But what is important is integrating all the elements together into a system. It must be practiced every day in a very consistent manner—not in spurts."

2.2.3 Lean Practices

The absence of a universal understanding of lean and LM makes it almost impossible to assess the impact of any development on LM as a whole. As a consequence, this research focuses on the more tangible lean practices as reference for assessing the impact of DBAs.

2.2.3.1 Definition and Selection

Lean practices are management practices that are generally associated with LM (Shah & Ward, 2003, p. 130). After LM has been widely established, several scholars have investigated and described essential elements of LM. Shah and Ward (2003, p. 131) have conducted an extensive literature review and provide a representative collection of lean practices. In total, they present 21 lean practices. However, only nine of the 21 practices have been found relevant by the majority of the reviewed publications.

To focus on the most relevant practices, this dissertation only considers those practices that have been stated by at least 50 percent of the reviewed literature. The result is a condensed list comprising nine lean practices. To ensure the actuality of the collection of lean practices, more recent literature on LM has been searched for new lean practices that need to be added to the selection. Thereby, value stream mapping was identified as an essential lean practice by many more recent publications (Albliwi, Antony, & Lim, 2015; Chiarini, Found, & Rich, 2016; Hines et al., 2004; Lander & Liker, 2007). Thus, value stream mapping was added to the list of highly relevant and established lean practices.

Table 6 provides an overview of the selected lean practices.

Table 6: Selection of key lean practices (adapted from Shah and Ward, 2003)

No.	Lean Practice	Category	Primary Objective
1	Preventive Maintenance	TPM	Maximize equipment availability by reducing unplanned maintenance and machine breakdowns.
2	(internal) Quality Management[11]	TQM	Ensure high quality by improving processes, identify defective products, and analyze the root cause of defects.
3	Continuous Flow Production	JIT	Enable continuous flow of a product through the value chain without interruptions and stockpiling.
4	Pull/Kanban	JIT	Prevent overproduction by triggering a process step only when the downstream process step needs replenishment.
5	Quick Changeover Techniques	JIT	Reduce the time for changing and setting up tools or machines.
6	Lot Size Reduction	JIT	Reduce the lot size to reduce inventory while increasing flexibility.
7	Value Stream Mapping*	JIT	Visualize the process to identify waste and opportunities for improvements.
8	Continuous Improvement	EMS	Set up and support a culture of active involvement in continuous improvement activities of all employees.
9	Cross-functional Work Force	EMS	Set up work teams with members of different backgrounds and a variety of skills to solve problems.
10	Self-directed Work Teams	EMS	Empower employees by setting up work teams with a high level of autonomy and individual responsibility.

* added to the selection. The practice is identified as a key lean practice by Albliwi et al., 2015; Chiarini et al., 2016; Hines et al., 2004; Lander & Liker, 2007; Womack & Jones, 2003.

[11] Shah and Ward (2003, p. 131) identified the lean practice "total quality management." Depending on the understanding of total quality management, this lean practice has a broad perspective including supplier and customer. As the scope of this thesis is the production facility exclusively, only the internal perspective of total quality management is considered. For the purpose of distinction, the lean practice is labeled "(internal) quality management."

2.2.3.2 Description

The Department for Production Management of the ITEM-HSG has designed a holistic approach to assess the operational performance of production systems. The assessment is part of the *St.Gallen Operational Excellence Benchmarking*. As the understanding of operational excellence has evolved from LM (Friedli & Bellm, 2013, p. 21), the performance assessment integrates most of the lean practices identified by Shah and Ward.

For this thesis, the selected lean practices are defined in accordance with the understanding of the St.Gallen operational performance assessment for two reasons. First, the assessment integrates several established lean concepts, such as the TPS, and is scientifically sound (Kickuth, Friedli, Loch, & Chick, 2006, p. 47). Second, it provides a clear structure by assigning lean practices to five distinct categories: total productive maintenance, total quality management, just-in-time, effective management system[12], and basic elements (Kickuth et al., 2006, pp. 48–52).

This structure is adopted for the description of the lean practices in the following section. While the St.Gallen operational performance assessment sets the basis for the understanding of the 10 selected lean practices, further literature is consulted for complementary details.

2.2.3.2.1 TOTAL PRODUCTIVE MAINTENANCE LEAN PRACTICES

The primary objective of *Total Productive Maintenance* (TPM) is to increase the efficiency of the production system by maximizing equipment availability. Therefore, TPM aims to set up a thorough maintenance system along the full equipment life cycle (Friedli & Bellm, 2013, p. 17). Besides maintenance, TPM includes housekeeping as well as continuous improvement and process upgrades through the integration of new technologies (Flynn, Schroeder, & Flynn, 1999, p. 250; Friedli & Bellm, 2013, p. 18). TPM encourages all employees to contribute to a high level of equipment quality and availability (Friedli & Bellm, 2013, p. 17).

Lean Practice 1: Preventive Maintenance

Preventive Maintenance is an important element of TPM and aims to minimize unplanned production interruptions due to unexpected machine breakdowns by performing maintenance tasks preventively (Friedli & Bellm, 2013, pp. 17–18). A reduction of unplanned machine stops increases the availability of the equipment, thus resulting in more stable production processes and higher productivity of the production system (Bertagnolli, 2018, p. 183).

[12] This category was initially named management system (Kickuth et al., 2006, p. 52) and updated later to effective management system (Friedli, Lembke, Schneider, and Gütter, 2013, p. 40).

2.2.3.2.2 *TOTAL QUALITY MANAGEMENT LEAN PRACTICES*

Total Quality Management (TQM) can be characterized as a "very rigorous problem-solving approach that is based on facts rather than on gut feeling" (Friedli & Bellm, 2013, p. 18). The primary objective of TQM is to identify and understand the causes for deviations within the production process, to eliminate the source of variation and thus to establish standardized and robust processes.

Lean Practice 2: (internal) Quality Management

Shah and Ward (2003, p. 131) identified *Total Quality Management* as key lean practices. The term can be understood in a broad sense, including supplier management, customer involvement, and cross-functional product development (Kickuth et al., 2006, pp. 49–50). As the unit of analysis of this dissertation is the production facility exclusively, the definition of the lean practice Quality Management follows a production-focused understanding. Quality Management in this understanding comprises quality inspection and failure root-cause analysis.

The purpose of quality inspection is to detect and sort out products that do not meet the required specifications. Arnheiter and Maleyeff (2005, p. 10) note that LM quality inspection prefers source inspection by employees and automated inspection over time-consuming end-of-line inspections. In the event of defective products, root-cause analysis methods, such as the DMAIC circle, are applied to gain a better understanding of the underlying factors of the defect (Friedli, Lembke, Schneider, & Gütter, 2013, p. 54).

2.2.3.2.3 *JUST-IN-TIME LEAN PRACTICES*

Just-in-Time (JIT) is a key element of the TPS. JIT lean practices allow companies to meet their customer requirements for a wide variety of product variants while keeping or even reducing their inventories (Friedli & Bellm, 2013, p. 19). The reduction of large inventory is desired for three reasons: first, inventory ties up capital; second, it takes up precious warehouse space; and third, it results in a comparably high number of defects (Holweg, 2007, p. 422). According to Ohno (1988, p. 75), JIT implies to "have all parts for assembly at the side of the line just in time for their user." Bertagnolli (2018, p. 84) describes JIT as a system to deliver the right material or service, in the right quantity and quality, to the right location and the right time. The idea of JIT is to produce only what is needed by the customer and therefore builds on a *pull* system.

Lean Practice 3: Continuous Flow

Continuous Flow Production describes a situation in which the material or product flows through the value-adding process steps without interruptions and waiting times in between (Friedli, Lembke et al., 2013, pp. 55–56; Womack & Jones, 2003, p. 21). In contrast to batch production, process steps in flow production are arranged in sequential order so that the product moves from one step to the next without being stocked. Therefore, flow production requires a shop floor layout that allows the installation of equipment in the right order. According to Bertagnolli (2018, p. 65), benefits of flow production are low inventory, low waiting times (and therefore minimal lead times), as well as high transparency of the material flow.

Continuous Flow Production requires a high level of equipment availability as a machine breakdown results in stopping the whole production line. Therefore, flow production is often accompanied by TPM practices (Womack & Jones, 2003, pp. 60–61). Furthermore, Continuous Flow Production requires the JIT provision of the right parts (Womack & Jones, 2003, p. 58). Finally, in Continuous Flow Production, it is crucial to ensure that products passed on to downstream processes are error-free, as otherwise the following process steps are forced to stop and wait for correct parts (Womack & Jones, 2003, pp. 60–61).

Lean Practice 4: Pull/Kanban

Pull/Kanban implies a demand-oriented material flow in which a process step only produces parts when triggered by the need for replenishment by the downstream process step (Womack & Jones, 2003, p. 67). Hopp and Spearman (2004, p. 142) define pull in contrast to push production as “a system, that explicitly limits the amount of work in process (WIP) that can be in the system.” They argue that Pull/Kanban has at least three benefits. First, Pull/Kanban limits the release of material in the process, thus reducing WIP levels. Second, by leveling the fluctuation of WIP levels, Pull/Kanban allows a smoother production flow. Third, quality is improved. As only limited stock is available in pull production, the system is pressured to improve quality to work without yield loss and rework.

The term *Kanban* is Japanese for *card* and describes the flow of information in a pull system. A Kanban card indicates when the minimal level of stock is reached. Based on the information on the Kanban card, an ordering process for replenishment from a warehouse is triggered. The Kanban system is used both internally and together with suppliers. However, Kanban also has two downsides. First, it is not applicable in cases of high fluctuations in lead times and demands. Second, due to the absence of security stock, a process step is highly dependent on the delivery availability of the upstream process step (Bertagnolli, 2018, pp. 86–87).

Lean Practice 5: Lot Size Reduction

The lean practice *Lot Size Reduction* is linked to the practice of *Quick Changeover Techniques* as the latter is the prerequisite of the former. Smaller lot sizes have a positive effect on the flexibility of a production plant (Friedli, Lembke et al., 2013, p. 55). Large lots, in contrary, result in high inventory before and after a process step (Bertagnolli, 2018, p. 31), thus causing high cost of tied capital as well as waste of storage space. Another benefit of small lot sizes is reduced rework due to insufficient quality as quality issues are detected faster and fewer parts are affected in comparison to large batches (Bertagnolli, 2018, p. 60).

Lean Practice 6: Quick Changeover Techniques

The primary purpose of the lean practice Quick Changeover Techniques is to reduce setup times between different products (Friedli, Lembke et al., 2013, p. 55). From a lean perspective, small lot sizes are desirable to increase flexibility. Small lot sizes, however, result in a high number of changeovers. Therefore, small lot sizes are only reasonable with short changeover times (Bertagnolli, 2018, p. 186). Quick changeovers have several positive effects: By reducing lot sizes, WIP inventory is reduced along with the cost of tied capital and required space. Also, flexibility is increased, thus allowing fast responses on order changes and shorter lead times. Finally, productivity increases because of less downtime due to changeovers (Bertagnolli, 2018, p. 189). To reduce changeover, Shigeo Shingo developed the SMED[13] method. The SMED method aims to reduce the changeover time to a single digit minute range by preparing and adjusting the new tool while the other one is still in action (Bertagnolli, 2018, p. 188).

Lean Practice 7: Value Stream Mapping

The lean practice *Value Stream Mapping* (VSM) is used to visualize the current status of a production process. VSM follows a holistic approach and integrates process steps along the whole value stream (e.g., from receiving the raw material until the delivery of the finished product to the customer). The identified value stream includes interfaces between different units along the value stream. VSM documents processing and waiting times, as well as the flow of material and the flow of information (Bertagnolli, 2018, p. 104; Friedli, Mänder, & Bellm, 2013, p. 309).

The objective of VSM is to analyze how lean the process is and to identify opportunities to reduce wasteful activities (Friedli, Lembke et al., 2013, p. 56). Due to the holistic end-to-end perspective of VSM, this practice helps to overcome local optimization but instead facilitates a process-oriented optimization approach. VSM not only helps to identify waste within the process but also its root cause. It is, therefore, an essential first step for end-to-end process optimization projects (Bertagnolli, 2018, p. 104).

2.2.3.2.4 EFFECTIVE MANAGEMENT SYSTEM LEAN PRACTICES

In addition to the three technical categories TPM, TQM, and JIT, the St.Gallen operational performance assessment comprises the category Effective Management

[13] Single-minute exchange of dies

System (EMS). A mature and well-designed EMS supports the management in "motivating and aligning people to work for a common goal" (Friedli & Bellm, 2013, p. 20).

According to Friedli and Bellm (2013, p. 20), the EMS has four main objectives: First, to ensure that employees can develop multiple skills that allow working in cross-functional teams. Second, to provide autonomy and a sense of belonging to the employees (e.g., by introducing working methods such as self-directed teams). Third, to involve employees in the continuous effort of process optimization and waste reduction. Fourth, to ensure the definition of consistent and challenging goals as well as ongoing management commitment.

Lean Practice 8: Continuous Improvement

Continuous Improvement (CI) originates from the Japanese philosophy of *Kaizen*. Kaizen is composed of the two Japanese terms *Kai* and *Zen* and translates to *change for the better*. It focuses on the continuous improvement of processes, services, and products. Kaizen is an integrative process (Bertagnolli, 2018, pp. 151–152). Key to Kaizen is the involvement of all employees into continuous thinking on improvements and waste reduction instead of leaving CI to dedicated industrial engineers. Ensuring employee involvement in CI activities is a major challenge for the management of manufacturing companies (Friedli & Bellm, 2013, p. 21). A well-designed suggestion program supports employees in providing improvement suggestions (Friedli, Lembke et al., 2013, p. 57). Most CI activities follow a rigorous scientific approach. Both well-known improvement cycles comprise a planning (PDCA) or an analyzing (DMAIC) phase that relies on accurate data from the process and requires a sound understanding of the underlying process (Sokovic, Pavletic, & Kern Pipan, 2010, p. 480).

Lean Practice 9: Cross-functional Work Force

The lean practice *Cross-functional Work Force* suggests to build teams consisting of employees with different backgrounds and a variety of different skills to improve productivity and quality by mastering challenging technical problems (Liker, 2004). Friedli, Lembke et al. (2013, p. 57) agree and characterize cross-functional project teams as an enabler for problem-solving. Setting up cross-functional work teams from different departments has several positive effects. On one hand, it creates a sense of unity of the workforce (Friedli, Basu et al., 2013, p. 110). On the other hand, according to Anand, Ward, Tatikonda, and Schilling (2009, p. 457), cross-functional project teams ensure a holistic perspective on process improvements.

Members of cross-functional teams need to have a range of skills related to teamwork, including the ability to listen to colleagues, to present their ideas, and to negotiate to identify the best solutions. Other critical competencies of team members are domain expertise as well as a sense of responsibility not only for their own part but for the overall result of the team's task (Chiarini et al., 2016, p. 250).

Lean Practice 10: Self-directed Work Teams

Self-directed Work Teams can be seen as a synonym for empowering employees (Shah & Ward, 2003, p. 134). Similar to the lean practice Cross-functional Work Force, the objective of Self-directed Work Teams is problem-solving. The practice, however, puts less emphasis on the diversity of background and skills of the group members but instead on the degree of autonomy of the team.

According to van Amelsvoort and Benders (1996, p. 160), Self-directed Work Teams are characterized by a maximum level of autonomy. For example, team members are responsible for their work assignment. Also, members enjoy a certain degree of freedom to select their work methods independently. Linked to a higher level of autonomy is a higher level of responsibility, thus employees should be empowered to make decisions without needing approval from the superior management level (Friedli, Lembke et al., 2013, p. 57).

2.2.4 The Role of Technology in Lean Manufacturing

The superior competitiveness of the TPS, the blueprint of the current LPS, does not originate from the extensive use of cutting-edge technology. Quite on the contrary, traditionally Toyota was not among the first companies that introduced new technology but rather lagged behind. Liker (2004, pp. 159–160) argues that this is not due to a general rejection of new technology but is rooted in the Toyota DNA of rigorously testing if new technology supports people and existing processes or not. Many technologies have failed this test, and therefore, Toyota rejected their introduction and relied on well-tried production equipment. Also, contrary to popular belief, LM does not reject automation per se. Already in the 1960s, Taiichi Ōno, one of the founders of the TPS, was in favor of process automatization with the process being supervised by employees (Kolberg & Zühlke, 2015, p. 1872). Although Toyota is not a leader in acquiring new technology, the company is still considered a global benchmark for using value-adding technology that supports core processes (Liker, 2004, p. 160).

At Toyota, new technology is seldom perceived as the savior for productivity problems, especially IT technology, which is viewed critically as it often increases the level of complexity. Bell and Orzen (2010, p. 53) emphasized that "the many benefits of electronic information systems are often offset by the waste they generate."

Maguire (2016) argues that lean and IT technology have been traditionally in conflict due to fundamental differences in the basic philosophy. For instance, while lean promotes simplicity, IT technology often introduces complexity. Another conflict is *push vs. pull*. Historically, production planning software planned the manufacturing process centrally, thus "'pushing' products through the manufacturing process"

(Maguire, 2016, p. 32). Lean, in contrast, advocates the pull approach, which is one of the five lean principles (Womack & Jones, 2003) as well as a key lean practice (Shah & Ward, 2003). Bell (2006, p. 16) agrees and lists “complexity versus simplicity, planning versus acting, one side pushing while the other is pulling” as ongoing conflicts between LM and information technology.

However, literature also presents contradicting results. Powell, Alfnes, Strandhagen, and Dreyer (2013, p. 324) argue that modern IT allows some conflicts to be bridged. For instance, recent developments in production planning systems enable a hybrid production control that combines the push and pull approaches. Other scholars present examples for IT support of existing lean practices. Kolberg and Zühlke (2015, p. 1871), for example, emphasize the potential of IT technology to support traditional Kanban systems. Coming back to Toyota, IT is critical to the company for many processes, such as financial transactions, handling of customer orders, and manufacturing process control. Like automation technology, IT technology needs to prove that it supports existing, value-creating processes in a test pilot before being introduced. Mikio Kitano, a plant manager of one of Toyota's largest production sites, is quoted by Liker (2004, p. 162): “At Toyota we do not make information systems. We make cars. Show me the process of making cars and how the information system supports that.”

In conclusion, it is a myth that LM is inherently resistant to apply leading-edge technology. However, LM is hesitant to accept new technology being pushed into the production system, without having demonstrated that it fulfills the critical criteria for new technology: to serve and support the people, process, and values of the LPS (Liker, 2004, p. 160).

2.3 Smart Manufacturing

This chapter consists of five parts. Chapter 2.3.1 discusses the origins of SM and provides a working definition. For a deeper understanding of the concept of SM, chapter 2.3.2 presents the *5C Architecture*. As SM is technology-driven (Metternich et al., 2017, p. 347), chapter 2.3.3 outlines core technologies that drive SM. Chapter 2.3.4 discusses the role of data utilization in SM, and chapter 2.3.5 introduces the manufacturing data lifecycle.

2.3.1 Origin and Definition

The term *smart manufacturing* was coined by the U.S. National Institute of Standards and Technology (2017). Accordingly, SM is a “fully-integrated and collaborative manufacturing system that responds in real-time to meet the changing demands and conditions in the factory, supply network, and customer needs.” Following Thoben et al. (2017, p. 7), Wallace and Riddick define SM as “a data-intensive application of information technology at the shop floor level and above to enable intelligent, efficient, and responsive operations.” Likewise, O’Donovan et al. (2015b, p. 2) consider SM as “the pursuit of data-driven manufacturing, where real-time data from sensors in the factory can be analyzed to inform decision-making.” The last definition provided by O’Donovan et al. is used as the working definition of SM in this dissertation.

SM is characterized as the fourth technical or industrial revolution. According to Lasi et al. (2014, p. 239), all industrial revolutions in the past have triggered technological leaps. Significant improvements in the field of mechanization were drivers of the first industrial revolution. The use of electrical energy caused the second industrial revolution, and computer technology triggered the third industrial revolution (Kang et al., 2016, p. 111; O’Donovan et al., 2015b, p. 5).

The objectives of SM are manifold. Thoben et al. (2017, p. 15) identified agility, quality, and efficiency improvements as well as economic and environmental sustainability as the most important objectives. Lasi et al. (2014, p. 239) consider individualization, flexibility, decentralization, and resource efficiency as major drivers for SM.

Research on SM is supported by governmental programs in the most important manufacturing countries. Under different terms[14], the United States, China, Germany, Japan, and Korea have formulated SM strategies. Although there are differences in the exact scope and focus, all strategies promote the use of modern IT technologies in manufacturing (Tao et al., 2018, p. 1; Thoben et al., 2017, pp. 5–8).

[14] In Germany the concept is called Industrie 4.0. China has initiated the China 2025 initiative. In the United States smart manufacturing and industrial internet are referring to the same basic concept (Tao et al., 2018, p. 1).

2.3.2 5C Architecture of Smart Manufacturing

For a better understanding of the concept of SM, the *5C Architecture* of SM proposed by Lee, Bagheri, and Kao (2015) is shown in Table 7.

Table 7: 5C Architecture for smart manufacturing (adapted from Lee et al. (2015, p. 19))

5C Architecture Level	Main Attributes	Main Function
V. Configuration Level	Self-configure and self-optimization	Decision implementation
IV. Cognition Level	Collaborative diagnostic and decision-making	Decision support
III. Cyber Level	Twin models for machines	Self-comparison
II. Conversion Level	Smart analytics	Self-awareness
I. Connection Level	Sensor network	Condition monitoring

Each level of the 5C architecture serves a different function (Lee et al., 2015, pp. 19–20; Lidong & Guanghui, 2016, p. 3; Qin, Liu, & Grosvenor, 2016, p. 175).

I. The *Connection level* is the foundation for the four higher levels by proving the hardware to collect, communicate, and store data from the manufacturing equipment. The data might be obtained from machine controllers, enterprise manufacturing systems, or directly by sensors.

II. By applying data analysis technologies, meaningful information is inferred on the *Conversion level* from the raw data collected on level 1. For instance, by predicting the remaining useful life, the conversion level increases the self-awareness of the equipment.

III. The *Cyber level* acts as a central information hub as it collects and processes data from every machine connected in the production system. A digital twin of the equipment allows cross-machine data analysis, such as peer-to-peer performance comparison.

IV. The *Cognition level* aims to collaborate and support manufacturing operators in decision-making. Therefore, the available data on individual machine status and comparative information has to be adequately visualized to transfer the acquired knowledge to the human counterpart.

V. The *Configuration level* is responsible for reconnecting the virtual cyberspace back to the physical world. It triggers corrective actions based on decisions taken in cyberspace to the machines in the physical world and therefore acts as a control system. Through feedback to the physical world, the production system applies self-configuration to increase resilience and adapt to variation.

2.3.3 Core Technologies

Based on a literature review on SM, Thoben et al. (2017) conclude that a variety of different definitions for SM exist. However, they also found that all descriptions of the concept consistently highlight advanced IT technologies as well as data analytics as critical elements to improve manufacturing operations (Thoben et al., 2017, p. 7). Similarly, Tao et al. (2018, p. 1) point out that all SM initiatives promote the integration of state-of-the-art technology in manufacturing systems. Hence, this chapter identifies and briefly defines four core technologies of SM.

Table 8 exhibits 10 technologies of SM according to twelve selected contributions to SM.[15] It indicates that the concept of SM comprises a wide range of technologies. Furthermore, it illustrates that there is no perfect consent among the reviewed scholars in regard to core SM technologies. However, the overview shows that four technologies are listed in almost all reviewed publications, namely *Big Data* (found in 10 of 12 publications), *Cloud Computing* (9), *Cyber-Physical Systems* (8), and *Internet of Things* (11).

Table 8: Selected technologies of SM and their appearance in key references

Technology	Key references											
	1	2	3	4	5	6	7	8	9	10	11	12
Additive Manufacturing						●			●			
Artificial Intelligence (AI) / Machine Learning (ML)		●					●			●		
Augmented Reality / Virtual Reality				●		●			●			
Big Data (incl. Data Mining)	●	●	●	●		●		●	●	●	●	●
Cloud Computing	●	●	●		●	●		●	●	●		●
Cyber-Physical (Production) System (CPS/CPPS)	●		●	●	●		●	●	●		●	
Internet of Things / Industrial Internet of Things	●	●	●		●	●	●	●	●	●	●	●
Robotics/Automation	●					●			●			
Simulation						●	●					
Smart Sensors (incl. RFID)			●	●				●				
References	1. Thoben et al. (2017), 2.Qi and Tao (2018), 3.Kang et al. (2016), 4. Prinz et al. (2018), 5.Sony (2018), 6. Lidong and Guanghui (2016), 7.O'Donovan et al. (2015b), 8.Pilloni (2018), 9.Y. Chen (2017)10.S. Wang, Wan, Li, and Zhang (2016), 11.Z. Li et al. (2017), 12.Lu (2017)											

15 Some papers address the concept of Industry 4.0. However, as Industry 4.0 and SM both refer to the same concept (Wuest et al. (2016, p. 23), SM and Industry 4.0 technologies are regarded as congruent and are used interchangeably.

The four outstanding technologies are considered as core technologies for SM. The remainder of this chapter introduces the four core technologies adding smart sensor technology, as it is an enabling technology for the four core technologies.

1. Smart Sensor Technology

The basic technology for collecting data in real-time are smart sensors. By collecting a range of different types of structured (times, quantities, process parameters) and unstructured (videos, picture, sound recordings) data, smart sensors are essential components of CPS and IoT networks and, therefore, fundamental for SM (Kang et al., 2016, p. 120). A key element of sensor technology is the *radio frequency identification device* (RFID) technology. RFID tags allow tracking the position of manufacturing objects. It is used for traceability of quality problems and to manage the material flow of the entire shop floor (Zhong, Huang, Dai, & Zhang, 2014, p. 828).

Today, data generated by embedded sensors in manufacturing exceeds the volume of two exabytes[16] annually, with higher rates expected in the future (S. Yin & Kaynak, 2015, p. 143).

2. Internet of Things (IoT)

The term IoT is frequently used in the literature, but there is no consensus about its definition (Tao, Cheng, Xu, Zhang, & Li, 2014, p. 1436). The definition below follows the recommendation of the *IoT Global Standards Initiative* (International Telecommunication Union, 2012, p. 7):

“Internet of Things: A global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies. Thing: With regard to the Internet of things, this is an object of the physical world (physical things) or the information world (virtual things), which is capable of being identified and integrated into communication networks.”

The IoT network collects data from smart sensors and exchanges data with information systems. Thus, the IoT network creates the database for big data analysis. The vision of IoT is to integrate the virtual with the real world by connecting all entities and facilitate communication by standard communication protocols. As IoT is not limited to applications within the manufacturing realm, the term *Industrial Internet of Things* (IIoT) for industrial assets such as machines, tools, and logistics operations, has emerged (Thoben et al., 2017, p. 8).

---

[16] 1 exabyte equals 10^{18} bytes or 1 billion gigabytes.

3. Cyber-Physical System (CPS)

CPSs are physical items (e.g., tools, machines, vehicles) that are equipped with sensors, microprocessors, communication interfaces, and actuators. CPSs are able to collect and process data of themselves and their environment (Thoben et al., 2017, p. 5). Also, they can communicate with other CPSs and interact with the physical world with actuators. Due to the communication interface with the internet or factory IT systems, CPSs have permanent access to real-time data of the factory. This enables a convergence of the physical world and the virtual world. The vision of SM contains a CPS that connects to other CPSs and forms decentral networks. Based on real-time data, these networks are supposed to optimize themselves autonomously (Vogel-Heuser, Bauernhansl, & Hompel, 2017, p. 11).

4. Cloud Computing

Cloud computing is essential for connecting physically dispersed production components. Cloud computing allows the central storage and processing of data from several sources. Also, it enables remote access to production equipment. However, although the user perceives the data as stored centrally, in fact, the data is stored on multiple servers at the same time. This guarantees a high level of accessibility to the cloud services and data, even in the case of a server breakdown (Vogel-Heuser et al., 2017, p. 135). Cloud computing provides scalable storage space and computational power. Thus, cloud computing meets the demand for big data analytics, without requiring large-scale investments in its own servers. Databases and analysis are accessible remotely through mobile devices connected to the cloud (Hagerty, 2017, p. 11). By enabling the integration of all devices that are equipped with communication technology in the same cloud, cloud computing provides the technological basis for IoT (S. Wang et al., 2016, p. 3).

5. Big Data

Big data generally refers to a set of data with a wide range, high volume, and complex structure that cannot be handled by traditional methods of data processing (Kang et al., 2016, p. 119). The objective of big data analytics is to extract from the tremendous amount of different data, meaningful and usable information, also called smart data (Kletti, 2015, p. 173). According to Kang et al. (2016, p. 119), big data analytics needs to provide effective analysis and visualization of data from different sources. The combination and analysis of data from various sources enable several improvement opportunities. Provost and Fawcett (2013a) argue that big data analytics, including *artificial intelligence* (AI) and *data mining* (DM), supports data-based decision-making. S. Wang et al. (2016, p. 6) add that big data analytics can provide real-time, complete, and effective information on every aspect of the factory, thus increasing the transparency of the manufacturing process.

2.3.4 Data in Smart Manufacturing

The relevance of data in SM is expressed by the two definitions of SM as “a data-intensive application of information technology at the shop floor level and above to enable intelligent, efficient, and responsive operations” (Thoben et al., 2017, p. 7) and SM as “the pursuit of data-driven manufacturing, where real-time data from sensors in the factory can be analyzed to inform decision-making” (O’Donovan et al., 2015b, p. 2).

As seen in the last chapter, SM implies a progressing integration of manufacturing technology, sensors, and IT technology. Thereby, the volume, accurateness, and real-time availability of manufacturing data increases (Tao et al., 2018, p. 2). As a consequence, the rate in which data is created in modern manufacturing systems shows significant growth (S. Yin & Kaynak, 2015, p. 143). However, not only has the volume of data dramatically increased, but also the ability to store and process data has significantly enhanced. According to Tao et al. (2018, p. 3), several cost-effective tools for data collection, storage, and processing have emerged recently.

SM strives to capitalize on both developments; that is, to apply new solutions for data collection and new techniques for data processing to convert raw manufacturing data into smart data to inform decisions. The potential positive impact of data exploitation is not limited to a particular use case. Data analytics can contribute to improvements in multifold aspects of manufacturing (O’Donovan et al., 2015b, p. 1).

Several research articles discuss use cases of exploiting manufacturing data. Zhang, Ren, Liu, and Si (2017) suggest big data analytics for optimizing maintenance processes. Chongwatpol (2015, p. 64) uses data analysis techniques, including DM, to lower defect rates, thereby improving the quality and overall performance of the manufacturing system. Similarly, Wuest, Irgens, and Thoben (2014) describe opportunities to improve quality by applying AI tools to data collected over the whole manufacturing process. According to Kusiak (2017), manufacturing companies become more and more aware of the strategic importance of data. Data—especially the ability to transform big data into smart data—will be a key factor for superior manufacturing competitiveness.

Data Analytics Capabilities

Following a publication of Gartner Inc. in 2016, many authors distinguish the following four levels of data analytics capabilities: *descriptive*, *diagnostic*, *predictive*, and *prescriptive* analytics (see Figure 5) (Banerjee, Bandyopadhyay, & Acharya, 2013; Hagerty, 2017; O’Donovan et al., 2015b; Shao, Shin, & Jain, 2014; Shuradze & Wagner, 2016).

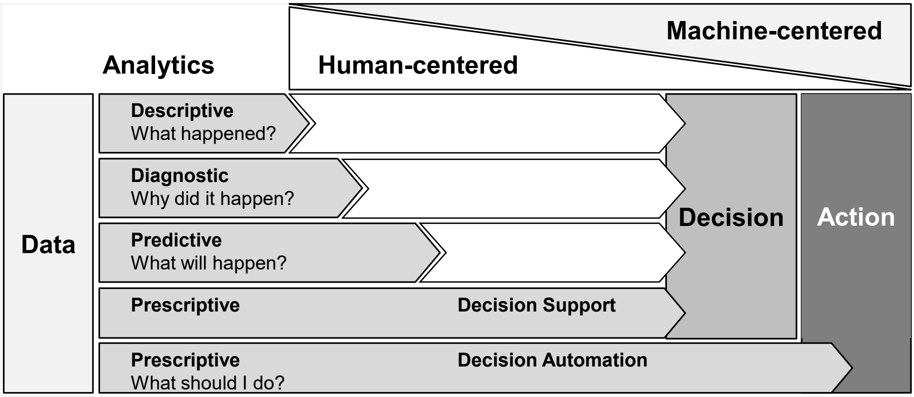

Figure 5: The Four Data Analytics Capabilities (based on Hagerty (2017, p. 10)

Descriptive analytics aims to answer the question of *what happened*. Descriptive analytics mainly provide different views of data, such as monitoring data from sensors and databases, to describe an observed phenomenon. At this first level of data analytics, no analysis of root causes is conducted (Banerjee et al., 2013, p. 2).

Diagnostic analytics addresses the question of *why did it happen*. Answering this question requires techniques like visualization to identify the root cause. Diagnostic analytics are valuable to understand the business environment better and are mostly used for strategic and forward-looking decision-making (Banerjee et al., 2013, p. 2).

Predictive analytics addresses the question of *what will happen* and seeks to predict potential future outcomes. By using statistical and DM techniques, drivers of observed phenomena are identified. In manufacturing, predictive analytics is applied to predict potential equipment breakdowns, thus enabling a more efficient maintenance system (Y. Chen, 2017, p. 592).

Prescriptive analytics addresses the question of *what should be done*. It combines describing, understanding, and predicting with suggesting approaches to achieve a desired future state. Prescriptive analytics compares various results of different decisions and presents the decision with the most desired expected outcome (Banerjee et al., 2013, p. 2).

Figure 5 also indicates a shift from human-centered decision-making to machine-centered decision-making. While the first three levels—*descriptive, diagnostic,* and *predictive*—merely provide the empirical foundation for informed decisions of humans, the fourth level, *prescriptive* analytics, already proposes an action. The next level towards machine-centered decision-making is *decision automation* (Hagerty, 2017, p. 10).

2.3.5 Manufacturing Data Lifecycle

As discussed in the previous section, data utilization is an enabler for SM. The most important and most challenging activity, thereby, is to convert big data from various sources into smart data. Smart data consists of meaningful information that can be utilized to inform concrete actions by the user (Kletti, 2015, p. 173; O'Donovan et al., 2015b, p. 3; Vogel-Heuser et al., 2017, p. 128). However, for the conversion from unstructured data to smart data, the data must undergo several steps. Tao et al. (2018, p. 4) refer to this process, including defining *data sources* (1), *data collection* (2), *storage* (3), *processing* and *analysis* (4), *visualization* (5), and *application* (6) as the *manufacturing data lifecycle*.

Figure 6 depicts the six steps of the manufacturing data lifecycle.

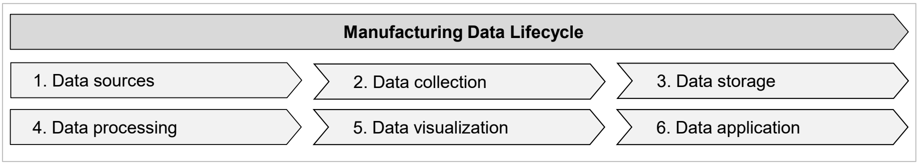

Figure 6: Manufacturing data lifecycle
(own illustration based on Tao, Qi, Liu, and Kusiak (2018) and Kletti (2015))

1. Data Sources

Tao et al. (2018, p. 3) present five sources of data, namely, *management data*, *equipment data*, *user data*, *product data,* and *public data*. Management data refers to all kind of data that is collected by manufacturing information systems (e.g., the manufacturing execution system MES or the enterprise resource planning system ERP). This type includes data related to product and production planning, material and inventory management, maintenance, as well as supply chain and financial data. Equipment data includes data such as real-time performance and operating conditions data. User data includes data from e-commerce providers (e.g., Amazon) and social networks (e.g., Facebook) and can be used to identify user preferences. Product data gathered by smart products with IoT functionality may contain product performance-related data (e.g., time) and environmental data (e.g., temperature).

2. Data Collection

Data originates from various sources and is collected primarily by sensors. This includes RFID technology that allows collecting product-related data such as processing time, environmental conditions, and current location (Saygin & Sarangapani, 2006). Built-in sensors in manufacturing equipment enable a permanent monitoring of the operational conditions (Tao et al., 2018, p. 4).

3. Data Storage

As the foundation for subsequent analysis, the collected manufacturing data needs to be stored and made available for all users (Tao et al., 2018). Cloud computing provides a convenient way to store structured and unstructured data. The scalability of cloud computing allows paying for just the right size of storage capacity and computational power needed. Also, it allows secure and remote access to data and services (Hagerty, 2017, p. 11).

4. Data Processing

Data processing describes activities to convert big data into actionable knowledge, often referred to as smart data. Smart data, in turn, is used to inform decisions (Kletti, 2015, p. 173). The processing step includes data cleaning and reduction, which involves activities to deal with redundant, missing, or inconsistent data. The cleaned data is then analyzed by various techniques. Gandomi and Haider (2015) have identified *clustering*, *association rules*, *regression*, *classification*, and *prediction analysis* as key data analysis methods. The type of analysis methods applied depends on the optimization problem and the structure of the data.

5. Data Visualization

The primary objective of data visualization is to support the understanding and interpretation of the presented data. Commonly used visualization techniques are various forms of charts and diagrams. Mittal, Khan, Romero, and Wuest (2017, p. 7) conducted a review on enabling factors of SM and found visual technology, such as *virtual reality* and *augmented reality*, as an important technology cluster.

6. Data Application

Data application are applications that use data to improve a certain aspect of the manufacturing process. Tao et al. (2018) distinguishes three categories of data applications: first *Design*; second *Manufacturing*; and third *Maintenance*, *Repair*, *and Operations*. Design can be improved by getting new insights into customer preferences and market trends. Applications in the manufacturing category comprise applications to track and monitor manufacturing equipment in real-time. MRO applications cover the manufactured products and include functionalities to identify product defects and the need for maintenance early, thus allowing precautionary actions.

2.4 Conjunction of Smart and Lean Manufacturing

LM and SM have often been treated separately in the past (Prinz et al., 2018, p. 21). Only recently, their relationship has received increased attention in OM research (Rossini et al., 2019, p. 2). According to Dombrowski, Richter, and Krenkel (2017), the existing literature discusses the conjunction of SM and LM from two perspectives: Either LM is considered as the basis for SM technologies, or SM technologies are seen as promoters of LM. Mayr et al. (2018, p. 623) added a third perspective, arguing that the combination of both manufacturing concepts yields positive synergies, without stating the direction of support.

Table 9 shows the perspective of 15 publications on the link between SM and LM.

Table 9: Perspectives on the conjunction of SM and LM

Perspective	**References**														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. LM is the foundation of SM	●	●	●	●							●				
2. SM advances LM					●	●	●	●	●	●	●	●			
3. SM and LM are mutually beneficial													●	●	●
References	1. Künzel (2016), 2. B. Wang et al. (2016) 3. Metternich et al. (2017), 4. Bick (2014), 5. Netland (2015), 6. Wagner et al. (2017), 7. Rafique, Ab Rahman, Saibani, Arsad, and Saadat (2016), 8. Rüttimann and Stöckli (2016), 9. Sanders et al. (2016) , 10. Riezebos and Klingenberg (2009), 11. Dombrowski et al. (2017), 12. Kolberg et al. (2016), 13. Tortorella and Fettermann (2018) 14. Rossini et al. (2019), 15. Lorenz, Buess, Macuvele, Friedli, and Netland (2019)														

1. Lean Manufacturing as Foundation for Smart Manufacturing

The perspective of LM as a basis for SM follows three different argumentations (Mayr et al., 2018, p. 623). First and found most often in literature, several authors argue that a successful introduction of SM requires robust, transparent, and standardized processes as a foundation (e.g., Dombrowski et al., 2017, p. 1063). The digitization or automation of inefficient processes will not reduce the inefficiency, but on the contrary, will intensify it. Second, authors such as Künzel (2016) promote lean as a critical foundation for SM, as lean principles force managers to always consider the impact of new technology on customer value and reduction of waste. Third, by reducing process and product complexity, LM facilitates the efficient application of SM technologies and tools (Bick, 2014, p. 46).

2. Smart Manufacturing advances Lean Manufacturing

The literature also provides support for this perspective. For instance, Rüttimann and Stöckli (2016) discuss the potential of CPPS as part of SM to increase the flexibility of existing LPS. Wagner et al. (2017, p. 130) analyzed the impact of SM technologies on lean elements such as 5S, Kaizen, and Pull, and conclude that industry 4.0 applications can support lean principles and have a positive impact on the stability of the system. Sanders et al. (2016, p. 827) developed a framework to identify barriers and challenges for lean implementation and evaluated how SM technologies can mitigate these barriers. They see great potential to overcome some of the existing shortcomings of LM through integrated information and communication (ICT) systems. Lean and ICT together can yield higher productivity and eliminate waste.

A promising concept called *Value Stream Mapping 4.0* is coined by Meudt, Metternich, and Abele (2017, p. 415) and summarized by Buer et al. (2018, p. 2930). They argue that traditional VSM is a manual *pen-and-paper* process and the necessary data collection time consuming and tedious. Furthermore, standard VSM only provides a snapshot of the current state of the process, not reflecting changes over time. By collecting real-time data, SM technologies reduce the effort and increase the accurateness and timeliness of data and thus enabling a real-time VSM system. Thereby, a dynamic picture of the shop floor is provided, which supports manufacturing employees in their decision-making.

3. Smart Manufacturing and Lean Manufacturing are mutually beneficial

The majority of contributions discussed above provide use case-based or exemplary evidence of support of LM to SM or vice versa. Quantitative studies investigating the relationship between LM and SM, however, are rather rare. One exception is Rossini et al. (2019) and Tortorella and Fettermann (2018). Rossini et al. (2019) conducted an empirical study with 108 European manufacturers and found a positive link between the implementation level of lean methods and a higher adoption level of industry 4.0 technologies. Similarly, Tortorella and Fettermann (2018) conducted a study on the relationship between LM and industry 4.0 among 110 Brazilian manufacturing companies and concluded that industry 4.0 and LM are positively associated.

A recent joint research paper of the University of St. Gallen and the ETH Zurich found supporting evidence from the Swiss manufacturing sector. As shown by Lorenz et al. (2019), the lean maturity of the company shows a significant positive correlation ($p<0.01$) with its digital maturity.

These quantitative studies indicate that there is no inherent contradiction between the human-centric lean approach and the implementation of SM technologies. However, they are not able to sufficiently explain how both concepts interact in concrete terms. This fact provides enough justification for further qualitative research to better understand the levers to increase performance by integrating both production paradigms.

2.5 Data Analytics - Methods and Techniques

The purpose of this chapter is to provide a basic understanding of the terms *data mining* and *machine learning* to the reader. For this introduction, the discussion of the mathematical basis of the presented methods and techniques is renounced.

The term *analytics* can be defined as “a scientific process of logical-mathematical transformation of data to improve decision-making” (Blum & Schuh, 2017, p. 258). The objective of data analytics is to concentrate and extract useful information hidden in a set of chaotic raw material (M. Chen, Mao, & Liu, 2014, p. 190). Furthermore, it seeks to identify inherent patterns to understand the logic behind the data and use this logic for extrapolation and forecasting. Although a variety of data analytics methods exist, this chapter focuses on the fundamental concepts of DM and ML. For a detailed introduction to a broad set of techniques related to data science for business, Provost and Fawcett (2013b) is recommended.

Data Mining

Following Choudhary, Harding, and Tiwari (2009, pp. 503–504), DM is an interdisciplinary field with the objective to uncover patterns in data and use this pattern to predict outcomes. DM includes techniques to find associations, anomalies, and patterns from a large set of data.

An overview of DM functions and techniques is shown in Table 10.

Table 10: Data mining functions and techniques (based on Choudhary et al. (2009, p. 515)

Data Mining Functions and Techniques				
Function	Association	Prediction	Cluster	Classification
Techniques	Decision tree	Association rule	Regression	Neural network
	Statistics	Fuzzy C means	Rough set	

The data sources may include historical data from databases as well as real-time data that are streamed into a system dynamically (Z. Li et al., 2017, p. 379). In the manufacturing context, description and prediction are the main objectives of DM. Descriptive DM seeks to describe the data set by identifying patterns and prescriptive

DM seeks to predict the future state of a model by extrapolating patterns of key variables found in the dataset.

Cross Industry Standard Process for Data Mining

To place a structure on the DM process, several guiding frameworks have been developed. Among these frameworks, the *Cross Industry Standard Process for Data Mining (CRISP-DM)* is one of the most widely applied ones (Harding et al., 2006, p. 970). The CRISP-DM shown in Figure 7 provides a high-level step-by-step instruction for DM.

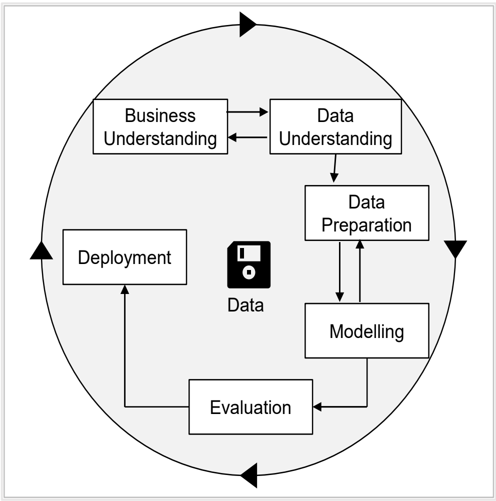

Figure 7: Cross Industry Standard Process for Data Mining (Provost & Fawcett, 2013b, p. 27)

It consists of six steps. The first two phases of *business understanding* and *data understanding* stress the importance of understanding the context of the intended analysis. There is no *one size fits all approach* for DM. One of the main challenges in DM is to transfer the business problem into a DM problem; thus, a good understanding of the business problem and context is crucial.

The *data preparation* phase is concerned with converting the data in appropriate forms to apply DM techniques. The actual application of DM techniques takes place in the *modeling* phase. In the *evaluation* phase, the DM results are assessed rigorously to ensure that they are valid and reliable before proceeding to the next step. In the *deployment* phase, the results of the DM are used to inform a decision. For instance, the DM process discovers patterns that are then used to implement a predictive model.

The process diagram visualizes the integrative approach. Going through the circle once without finding the expected solution is rather the rule than the exception. However, every iteration increases the understanding of the DM team of the underlying problem and the available data (Provost & Fawcett, 2013b, pp. 26–32).

Machine Learning

"ML is allowing computers to solve problems without being specifically programmed to do so" (Wuest et al., 2016, p. 25). This understanding goes back to Samuel (1959), who programmed a computer to learn to play checkers. In the context of manufacturing, ML techniques may be used to identify complex patterns within raw data and, based on this underlying logic, derive models. These models are applied for forecasting, classification, and prediction. From the basis of ML, the concept has developed over time, resulting in a variety of different techniques, algorithms, and application areas. However, no shared understanding of the structure of the different element exists. For an overview of the main ML techniques, Figure 8 presents a classification following the work of Pham and Afify (2005, p. 401). They also provide a brief review of the shown ML techniques.

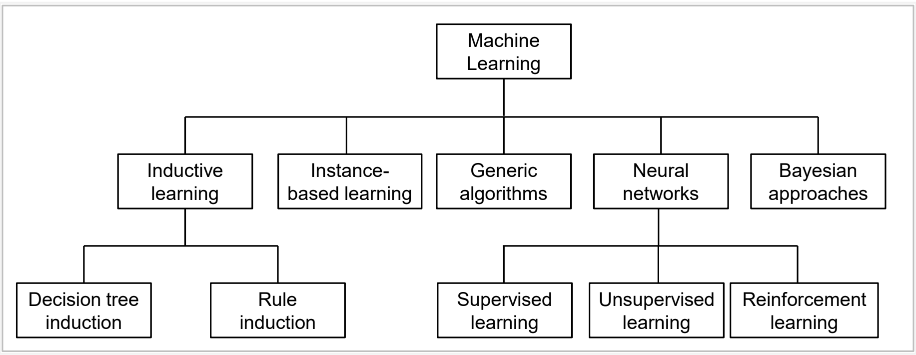

Figure 8: Classification of machine learning techniques (Pham & Afify, 2005, p. 401)

Artificial Intelligence

Besides ML the term AI (artificial intelligence) is often used in the literature. There are different voices on the question of whether AI and ML are the same, or one is part of the other. For instance, Wuest et al. (2016, p. 24) describe AI as a subgroup of ML. Quite on the contrary, Rimpault, Balazinski, and Chatelain (2018, p. 1) see ML as part of the AI domain, and Z. Li et al. (2017, p. 381) describe AI and ML as distinct parts of DM. In this dissertation, both terms will be used interchangeably and refer to the ability of computers to learn from databases and find hidden patterns in the data.

Wan, Yang, Wang, and Hua (2018) conducted a comprehensive review of AI applications in smart factories and link application objects with the respective AI method and resulting improvements. In total, they list 18 application objects for AI in a smart factory.

Contrasting DM and ML, two main differences can be found: first, ML can perform self-learning by adapting the identified rules as per the scenario. DM, on the contrary, follows predefined rules. Second, DM requires human involvement all the time, while in ML humans are only involved a the initial definition of the ML algorithm. Afterward, it will conclude everything by its own means. In conclusion, DM and ML are techniques used to analyze a large set of chaotic data and, therefore, fit the conditions often found in manufacturing. Especially in the light of the advancing integration of sensors and advanced ICT technology in manufacturing, the volume of accurateness of manufacturing data will dramatically increase (Tao et al., 2018, p. 1). Consequently, the application of ML and DM is becoming increasingly relevant to exploit the data to optimize the production system.

3 Smart Manufacturing Technologies in LM – Quantitative Study Results

This chapter presents selected results from a quantitative survey on the future of LM conducted by the author in 2017. Chapter 3.1 introduces the concept and objectives of the study, as well as the data collection process. Chapter 3.2 presents general findings and highlights differences between mature lean companies and the overall sample. Chapter 3.3 summarizes the findings and derives implication for the following chapters.

3.1 Introduction and Data Collection

3.1.1 Introduction

Inspired by the success of LM in the automotive industries, other industries within and outside the manufacturing sector have adopted many elements of LM. The focus on customer value, elimination of waste, and the principles of *Pull* and *Flow* are today found in many companies of different sizes and products. The high relevance of lean in practice was the essential motivation to assess the status quo of lean 30 years after the term was coined in 1988. Furthermore, after a long time of adopting and refining existing lean practices from the TPS, the question arose: Quo vadis lean?

Therefore, the objective of the study was to assess the current state of lean across different industries and to take a look in the future. To reflect this intention, the title of the study was: *Lean2020 – The Future of Operational Excellence.*

To ensure the practical relevance of the results, the study was designed in cooperation with three international manufacturing companies from the aviation, pharmaceutical, and mechanical engineering sectors. The study was carried out following an established procedure of a benchmarking study, comprising four phases. In the first phase, the questions are derived from current academic discussions on lean as well as from input of the benchmarking industry partner. The questions are clustered, and a questionnaire developed and tested. In the second phase, the questionnaire is implemented as an online survey and sent to a set of companies. The raw data from the online survey is evaluated and, based on the maturity level of participating companies, a condensed sample of so-called *Leading Companies* is built. The maturity level is calculated by ex-ante defined metrics, such as the level of implementation of various lean tools and the spread of lean within the company.

In the third phase, interviews with the Leading Companies are scheduled to verify the information given in the only survey and to collect additional context information on the company and their lean approach. Based on these interviews, the research team selects the 10 most mature companies and creates anonymous case studies about the companies. Based on the case studies the industry partners select five companies which are, from their perspective, most promising to learn from.

These five companies are called *Successful Practice* (SP) companies. To get real-world insights and enable a face-to-face discussion, these companies are visited in one-day site visits during the fourth and final phase of the benchmarking procedure.

3.1.2 Data Collection

The online survey was sent to more than 500 companies, mainly in the German-speaking area. Contacts have been lean managers from global lean units as well as lean managers at a production site, plant managers, and production managers. The online survey took place in the second and third quarter of 2017. From all questionnaires filled, 75 were qualified for subsequent analysis.

Sample Structure

The structure of the study sample is shown in Table 11.

Table 11: Sample structure of study: Lean2020 – The Future of Operational Excellence

Sample Structure Study: Lean2020 – The Future of Operational Excellence			
Participants	75	**Industries**	14
Revenue	51% of all participants generate more than 250 million EUR in revenue	49% of all participants generate less than 250 million EUR in revenue	
Employees	74% large companies with more than 250 employees	26% small and medium-sized enterprises (SMEs) with less than 250 employees	
Lean Experience	50% of all participants have five or more years of Lean experience	25% of all participants have even more than 10 years of Lean experience	
Company Type[17]	68% of all participants are in Industrial Goods (B2B) business	31% of all participants are in Consumer Goods (B2C) business	

[17] One percent of the participating companies is working primarily for the government (B2G)

3.2 Study Results and Successful Practice Companies Patterns

As discussed in chapter 2.2.1, there are a plethora of lean definitions in the academic world. The same applies to the industry. Therefore, the term lean may be used slightly differently within the participating companies. The lean understanding of the study follows the five lean principles that *define value from the customer perspective* (1), *identify the value stream* (2), *flow* (3), *pull* (4), and *strive for perfection* (5), and includes lean outside the manufacturing realm, such as logistics but also R&D and service.

In the remainder of this chapter, selected results of the study are presented and briefly discussed. They have built the foundation and motivation for this research. Some of the results have been published in previous or amended versions in the following publications.

- *Lorenz, Rafael; Buess, Paul; Macuvele, Julian; Friedli, Thomas; Netland, Torbjørn H. (2019): Lean and Digitalization—Contradictions or Complements? In: Farhad Ameri, Kathryn E. Stecke, Gregor von Cieminski und Dimitris Kiritsis (Hg.): Advances in production management systems. Production management for the factory of the future: IFIP WG 5.7 International Conference, APMS 2019, Austin, TX, USA, September 1-5, 2019: proceedings, Bd. 566. Cham: Springer (IFIP Advances in Information and Communication Technology, 566), S. 77–84.*
- *Macuvele, J., Buess, P., Friedli, T. (2018): Lean2020 – The Future of Operational Excellence. Final Report. Institute of Technology Management at the University of St. Gallen, St.Gallen.*

Lean Remains Critical for Future Competitiveness

As shown in Figure 9, almost all participants believe that the relevance of lean for competitiveness will either increase or remain as it is today. Only a tiny fraction expects a decreasing relevance of lean. This assessment applies to the overall sample and the five successful practice (SP) companies.

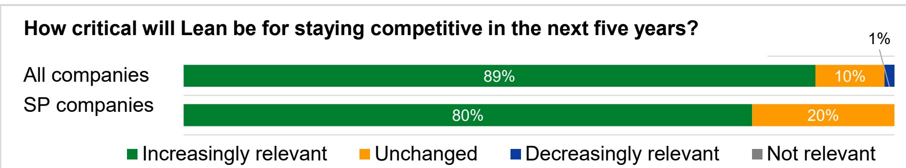

Figure 9: Lean and competitiveness

Digitalization will Support Lean

Especially interesting for the topic of this dissertation is the perspective of how lean companies deal with emerging digital technologies. Like academic literature, the overall sample takes two different perspectives. The majority of 56 percent assumes that digitalization will advance lean (see Figure 10). Thirty-eight percent also expect a positive interrelation between lean and digitalization but consider the former as the foundation of the latter. A coexistence of lean and digitalization without major mutual impact is predicted by only four percent. The trend is in line with the expectations of the SP companies. Four out of five see digitalization as a promising way to enhance lean.

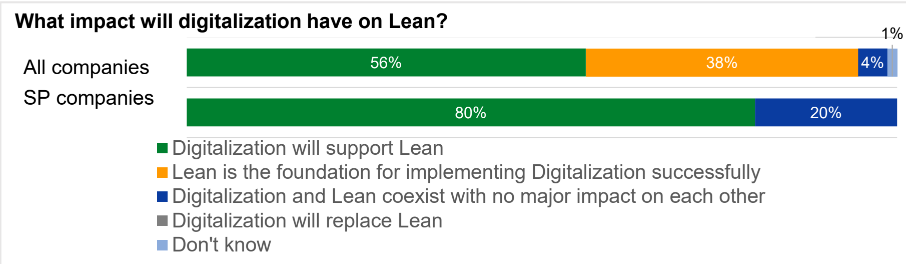

Figure 10: Impact of digitalization on lean

Interestingly, no participant expects a complete shift from lean to a new production paradigm driven by digital technologies.

Big Data is the Digitalization Trend with the Highest Potential to Support Lean

In the study, participants are asked to evaluate the potential of four high-level technology trends to support lean. According to Figure 11, big data collection and analytics is the technology trend with the highest potential to support lean.

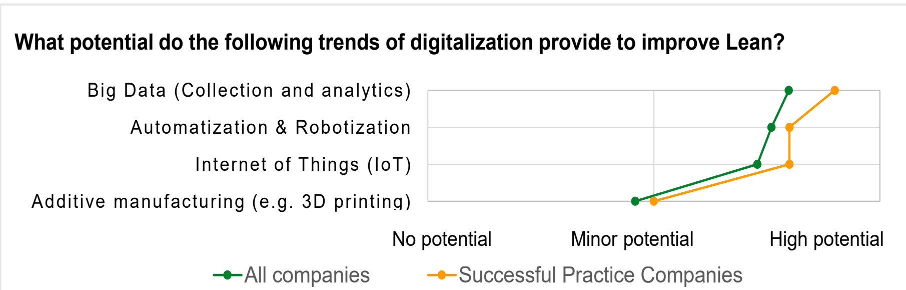

Figure 11: Potential to improve lean of selected trends of digitalization

While on average, a medium potential is also foreseen for the technologies of automatization and robotization as well as IoT, comparably few expect an important role of additive manufacturing to enhance lean.

Following Figure 11, the SP company group is more optimistic in general and especially considers big data as a promising technology to take lean to a new level. In response to the question, which of these technologies can support the five lean principles, big data was also selected more often than the other trends—especially regarding defining customer value, identifying the value stream, and seeking perfection—and participants consider big data as a promising technology trend.

Given the expected relevance of Big Data, Figure 13 addresses the share of equipment with real-time monitoring, while Figure 14 depicts the maturity of data analytics in several areas.

Lack of Management Capabilities and Shortage of Manpower are Barriers to Use Digital Technologies to Improve Lean

As this dissertation is interested in identifying key enablers of DBAs, Figure 12 is a good starting point to look at. It provides a list of potential barriers companies face when planning to utilize digital technologies to support their LPS.

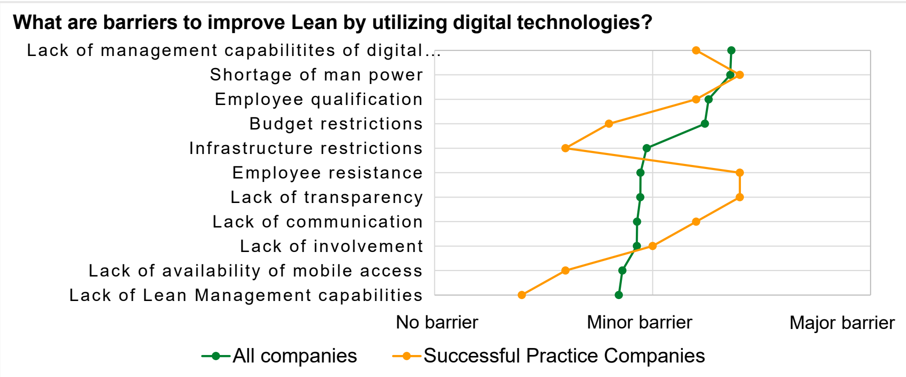

Figure 12: Barriers to improve lean by utilizing digital technologies

From an overall perspective, the main barrier is the *lack of experience and capabilities* to manage the introduction of digital technologies. Furthermore, participants report a *shortage of manpower* especially of employees with the right *qualification*. Financial restrictions are also perceived as an important barrier. The remaining elements on the list are considered comparably less challenging. Some elements that have been critical to the implementation of lean, such as employee acceptance and involvement, communication, and transparency, are still rated on average as a minor barrier. In contrast to the lack of skills to manage digital technologies, most companies are sure to have the right LM capabilities in place.

The evaluation of barriers of the SP group is quite diverse from the overall sample. For this group, aspects such as *employee resistance*, *lack of communication*, and *involvement* are much more prevailing than financial or infrastructure restrictions. It appears that the SP companies have learned the importance of employee acceptance and involvement for new ways of working from their lean journey and, as a consequence, emphasize these soft factors.

Successful Practice Companies Monitor More Equipment in Real-time

The positive expectation towards data analytics of SP companies seems to be reflected in the investment to monitor production equipment to collect data (see Figure 13).

While within the overall sample only 13 percent have equipped more than 80 percent of the production equipment, and only 30 percent more than 40 percent, the proportion among the SP companies is 60 percent and 80 percent, respectively.

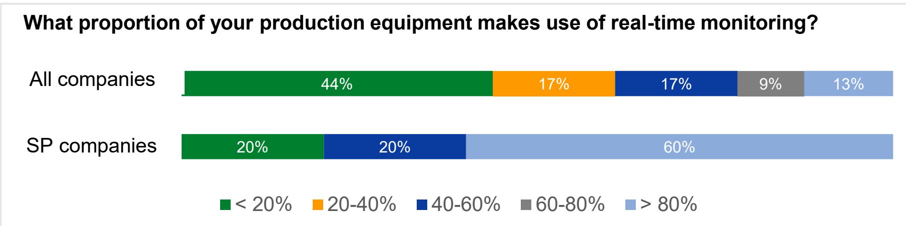

Figure 13: Real-time monitoring of production equipment

Data Analytics is on Average on the Maturity Level Diagnostic

Figure 14 illustrates the current status of data analytics maturity regarding the four levels of data analytics capabilities introduced in chapter 2.3.4.

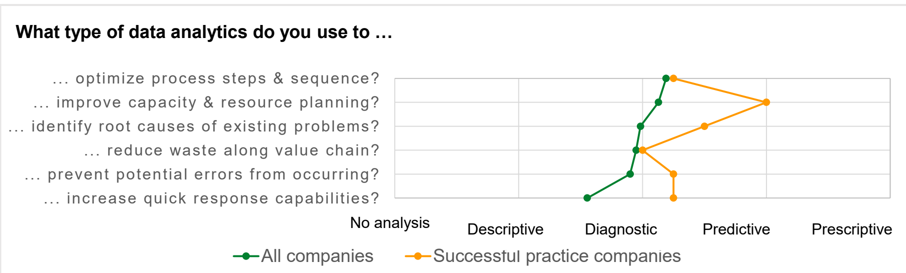

Figure 14: Applied type of data analytics

In general, participants report that they are currently operating on the diagnostic level to optimize processes, improve planning, identify root causes, reduce waste along the value chain, and prevent failures from reoccurring. SP companies not only collect more data due to a higher level of production equipment monitoring, but also utilize their data on a high analytics level. For capacity and resource planning, these companies have already reached the predictive analytics level. A deep dive in the survey data reveals that prescriptive analytics is currently the exception.

3.3 Implications

The study suggests that the time of lean is far from over. Quite on the contrary, a substantial majority of study participants expect even a rising relevance of lean for the future competitiveness of their companies. With regard to the interplay of lean and digitalization, the study results are in line with the current discussion of academic literature. Both perspectives, lean as the foundation of digitalization and digitalization as a means to enhance lean, are represented with a majority supporting the second perspective. From a set of four digital technology trends, big data was selected as most promising to support lean. This applies to the whole group, especially to SP companies. These companies are also pioneering in data collection (high share of real-time monitoring) and data-analytics (higher analytics maturity level). These findings support the underlying assumption of this thesis, that data analytics will have a positive impact on the manufacturing industry, including companies operating an LPS.

Although valuable to get a quantitatively backed overview of the role of data in manufacturing and its potential for LM, the study is not capable of addressing the questions of *how* data utilization can support LM. The research of this thesis is motivated by the aspiration to close this gap. Therefore, the following chapters intend to identify actual applications of data utilization in manufacturing and investigate challenges and enablers in more detail.

The following, chapter 4, starts with a comprehensive literature review to identify common DBAs and summarizes inherent challenges. Furthermore, based on literature input, theoretical reasoning, and discussions with researchers and practitioners, propositions on the impact of DBAs on lean practices are derived and visualized in a DBA – Lean Practice Impact Matrix. Chapter 5 complements these findings with qualitative case study and expert interviews, focusing on the identification of key challenges and enablers for applying DBAs.

4 Data-based Applications in Lean Manufacturing

This chapter has three objectives related to the three SRQ. Based on a comprehensive literature review, chapter 4.1 addresses the first SRQ by identifying use cases of DBAs in the manufacturing industry. Furthermore, it summarizes key requirements described in the DBA use cases, thus contributing to answering the SRQ 2. Afterward, SRQ 3 is addressed in chapter 4.2, which discusses the potential implications of DBAs for established lean practices

4.1 Data-based Applications

This chapter comprises seven subchapters. Chapter 4.1.1 provides a definition of the term *data-based application*. Chapter 4.1.2 describes the process of identifying DBA use cases from the literature. Chapter 4.1.3 provides an overview of identified use cases and proposes a classification system. The individual DBA use cases are presented in chapter 4.1.4. Shared core functions of DBAs are identified and discussed in chapter 4.1.5. Chapter 4.1.6 summarizes the key requirements found in the DBA use cases, and chapter 4.1.7 closes with a summary.

4.1.1 Definition: Data-based Applications

The term *data-based application* (DBA) is a word creation of this dissertation and as such not directly adopted from existing publications. The term *application,* however, is frequently used in the context of data analytics. For instance, Russom (2015, p. 2) uses the term *datadriven applications* to describe IT applications that automate "business processes, problems, and opportunities that can only be driven forward, solved, and leveraged via ample volumes of diverse data."

Other scholars refer to the term *application* in various ways, without providing a specific definition. Åkerman et al. (2018, p. 411) use the term *Big Data applications* in the context of enabling data-driven decisions. Wuest et al. (2016, p. 25) use the term *machine learning applications,* Mehta, Butkewitsch-Choze, and Seaman (2018, p. 1043) the term *manufacturing analytics application*, and Shao et al. (2014, p. 2192) the term *DA* (data analytics) *application.*

Although the term application is not explicitly defined, is it reasonable to assume that all quoted scholars share the basic understanding of the term as Tao et al. (p. 5). They have suggested a six-step manufacturing data life cycle (see chapter 2.3.5) comprising the steps data sources (1), data collection (2), data storage (3), data processing (4), result visualization (5), and data application (6). This dissertation follows this understanding of a *data application as the final output of the manufacturing data lifecycle.*

The term DBA is defined in this dissertation as an umbrella term for use cases of data utilization in manufacturing. To be classified as a DBA, an application must meet the following four criteria:

1. According to the Cambridge University Press (2019a), an application is “a way in which something can be used for a particular purpose.” In line with this understanding, a DBA needs to serve a particular purpose. Therefore, data collecting and storing data for the sake of having the data available is not a DBA.
2. A DBA is dependent on the availability of data. Therefore, applications such as repetitive automatization are not considered as a DBA.
3. The data collection and data analysis are enabled by emerging (smart) technologies (e.g., smart sensors). Thus, traditional six sigma, including manual data collection, is not considered as a DBA.
4. DBAs considered in this dissertation are applicable in a manufacturing facility. Although literature has shown the potential of DBAs to increase performance over the whole value chain,e.g., by predicting customer demand (Harding et al., 2006, p. 974) or integrating suppliers (Kolberg & Zühlke, 2015, p. 1871), this dissertation does not consider applications that go beyond the boundaries of a factory.

4.1.2 Search Process

The process to identify relevant use cases of DBAs follows the literature review process introduced in chapter 2.1 and includes the same five online databases. The applied search terms and search operators are shown in Table 12.

Table 12: Applied search terms and operators for DBA literature review

<table>
<tr><th>ST 1</th><th>OP 1</th><th>ST 2</th><th>OP 2</th><th>ST 3</th><th>OP 3</th><th>ST 4.1</th></tr>
<tr><td>Smart
Digital</td><td>AND</td><td>factory
manufactur*
technology</td><td>AND</td><td rowspan="3">data
“ ”</td><td rowspan="3">AND</td><td>application
example
use case</td></tr>
<tr><td rowspan="2">Industry
4.0
Industrie
4.0</td><td rowspan="2" colspan="3">AND</td><td>ST 4.2
Analytics/analysis
maintenance
fault detection
root cause</td></tr>
<tr><td>ST 4.3
challenges
requirements
capabilities</td></tr>
<tr><td colspan="7">ST: search term, OP: operator</td></tr>
</table>

The selected search terms cover the combination of all terms related to SM (smart manufacturing, industry 4.0, etc.) and data. The keywords under the search-term (ST) 4.1 were selected with the purpose to find use cases and examples of data utilization in manufacturing. The keywords under ST 4.2 were selected due to prior knowledge that functions such as predictive maintenance might be realized with DBAs. The third group of keywords, ST3, intends to identify papers that deal with challenges and capabilities of data utilization in manufacturing in general, thus providing first insights to answer SRQ 2.

As in the literature review in chapter 2, this review includes journal papers as well as papers of renowned conferences. Conference papers not only tend to be more recent but also focus more often on actual use cases and provide context to these use cases whereby journal papers rather conflate the main findings of several use cases. After the initial keyword search, forward and backward search was applied to identify further relevant articles. The result of the search process is shown in Table 13 to Table 15.

4.1.3 Classification and Overview

The DBA classification system presented in Figure 15 is inspired by several existing classifications in the literature. These classifications mostly comprise applications of a particular type, such as *data mining* applications or *artificial intelligence* applications.

Table 13 presents five classification systems of applications with four different classification objects related to data utilization in manufacturing. These applications should be covered by a comprehensive collection of DBAs in manufacturing and serve as the basis for the overview of DBAs shown in Figure 15.

Table 13: Classifications of applications related to data utilization

Classification object	**Data-driven methods**	**Data mining application**	**Data mining application**	**Artificial intelligence application**	**Data analytics enabled I.40 application**
Applications	Process and planning*	Design	Quality control*	Design	Asset utilization*
	Business and enterprise	Manufacturing systems*	Job shop scheduling*	Process planning*	Quality control*
	Maintenance and diagnosis*	Shop floor control and layout*	Fault diagnostic*	Quality*	Supply Chain Management
	Supply chain	Fault detection & quality improvement*	Manufacturing process*	Maintenance and diagnosis*	Product Monitoring*
	Transport and logistics*	Maintenance*	Maintenance*	Control*	Workplace Safety*
	EHS*		Defect analysis*	Scheduling*	
	Product design		Yield improvement*		
	Quality management*		Condition-based monitoring*		
			CRM		
Reference	O'Donovan, Leahy, Bruton, and O'Sullivan (2015a, p. 8)	Harding et al. (2006, p. 972)	Choudhary et al. (2009, 504 - 513)	Meziane et al. (2000, p. 18)	Pilloni (2018, p. 5)

* Applications are applicable in manufacturing. EHS: Environment, Health, and Safety

Instead of clustering the DBAs according to the underlying technique (e.g., data mining), the classification in Figure 15 is based on their primary purpose. It comprises six DBA categories: Planning and Scheduling (I); Production Control (II); Maintenance (III); Internal Logistics (IV); Product Quality Management (V); and Environment, Health, and Safety (VI).

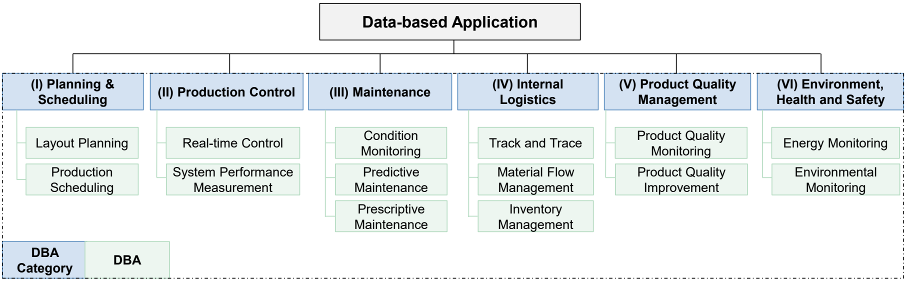

Figure 15: Overview and classification of data-based applications

Each DBA category comprises two to three DBAs. DBAs belonging to the same DBA category have similar objectives but might follow a different approach to achieve them. For instance, the DBA category *Product Quality Management* aims to ensure a high quality of the finished product. The two DBAs of this category support this objective but in different ways. The DBA *Product Quality Monitoring* detects faulty parts and sorts them out, thus preventing the shipment of defective products to the customer. The DBA *Product Quality Improvement* uses historical data on product defects to identify its root cause. Thereby, this application systematically reduces the risk of defective products.

The presented classification system of Figure 15 covers all manufacturing-related applications listed in Table 13. It also integrates applications of data utilization that have not been included in the reviewed classification systems but have been identified during the DBA use case search process. In total, the classification system consists of six categories, comprising 14 DBAs. Although all categories and DBAs presented in Figure 15 originate from literature, the titles are no 1:1 replication of terms from literature but instead reflect a synthesis of different terms used in the literature. The result of the literature review for DBAs is presented in Table 14 and Table 15. Table 14 depicts which DBAs are listed in review articles or general articles on data utilization. Only peer-reviewed journal publications have been considered for this overview. Table 15 then provides an overview, including a short description of each DBA.

Table 14: Data-based applications in key references

	Scope	Big Data			DM		ML		IC	SM / I4.0			
Category	**Data-based Application**	1	2	3	4	5	6	7	8	9	10	11	12
I. Planning and Scheduling	Production Scheduling	●	●	●	●	●	●	●	●	●			●
	Layout Planning		●	●	●	●							
II. Production Control	Real-time Control		●		●	●			●	●	●	●	●
	System Performance Measurement			●									●
III. Maintenance	Condition Monitoring	●	●	●		●	●			●	●	●	●
	Predictive Maintenance	●	●	●		●	●	●		●	●	●	●
	Prescriptive Maintenance												
IV. Internal Logistics	Track and Trace	●				●				●	●	●	
	Material Flow Management	●								●	●	●	
	Inventory Management	●			●			●					
V. Product Quality Management	Product Quality Monitoring	●	●		●	●		●		●	●	●	●
	Product Quality Improvement		●	●	●	●	●			●	●	●	●
VI. Environment, Health, and Safety	Energy Monitoring	●	●	●						●	●	●	●
	Environmental Monitoring		●								●		
References	1. J. Li, Tao, Cheng, and Zhao (2015), 2. O'Donovan et al. (2015a), 3. Yan, Meng, Lu, and Li (2017), 4. Choudhary et al. (2009), 5. Harding et al. (2006), 6. Wuest et al. (2016), 7. Wan et al. (2018), 8. Zhang et al. (2014), 9. Tao et al. (2018), 10. Pilloni (2018), 11. Lidong and Guanghui (2016), 12. Illa and Padhi (2018)												
Key	DM: data mining, ML: machine learning; IC: information capturing; I4.0: industry 4.0												

The twelve publications of Table 14 originate from different literature streams. References 1–3 primarily address the topic of *Big Data,* for instance, in PLM (Reference 1), in manufacturing (2), and in an Industry 4.0 environment (3). References 4–5 are concerned with DM in manufacturing. Referencse 6–7 address applications of ML/AI in manufacturing in general (6) and in a smart factory setting (7). Reference 8 takes a different perspective and investigates real-time information captured in an IoT environment. Finally, references 9–12 are considering data utilization as part of SM and deal with data-driven SM (9), data collection and Big Data as pillars of industry 4.0 (10), Big Data as part of CPSs in industry 4.0 (11), and the transition to smart factory with Big Data and edge analytics (12).

Table 15 presents a short description of every DBA, including its objective (O) and a possible approach (A).

Table 15: Data-based applications short description and use case references

C	Application	DBA Short Description	R	P
I	Production Scheduling	O: Determination of an optimal production plan. A: Mathematical optimization model using data (e.g., machine availability) to optimize a target value (e.g., cycle times).	1 2	67
	Layout Planning	O: Determination of an optimal production layout. A: Mathematical model optimizing a target value (e.g., minimal WIP, minimal material handling) under given constraints (e.g., number and type of machines, space, flow orientation).	3 4	69
II	Real-time Control	O: Real-time monitoring of the production process. A: Permanent comparison between the expected behavior and the actual behavior of the production system, including real-time notification in case of deviations.	5	70
	System Performance Measurement	O: Transparent overview of the overall system performance and visualization of trends. Use of metrics for benchmarking. A: Automatic collection of operations metrics (e.g., times, scrap rates) and calculation and visualization of KPIs (e.g., OEE).	6 7	71
III	Condition Monitoring	O: Increase maintenance effectiveness and efficiency. A: Monitoring equipment health status to trigger maintenance activities only in the case of unusual behavior.	8	73
	Predictive Maintenance	O: Increase maintenance effectiveness and efficiency. A: Equipment condition monitoring and prediction of degradation to derive maintenance plans that ensure equipment availability (effective) but avoid unnecessary maintenance (efficient).	9 10	74
	Prescriptive Maintenance	O: Increase maintenance effectiveness and efficiency. A: As Predictive Maintenance but in addition, the Prescriptive Maintenance application also suggests, or even starts, maintenance activities autonomously.	11	75
References		1. Vallhagen, Almgren, and Thörnblad (2017) (C), 2. Zhong et al. (2014), 3. Kumar, Singh, and Lamba (2018), 4. Agard and Cunha (2007), 5. Hirmer et al. (2017), 6. Meissner, Müller, Hermann, and Metternich (2018) (C), 7. Friedemann, Trapp, Stoldt, Langer, and Putz (2016) (C), 8. Yunusa-kaltungo and Sinha (2017), 9 . Z. Li et al. (2017), 10. Åkerman et al. (2018), 11. Matyas, Nemeth, Kovacs, and Glawar (2017)		

C: Category, O: Objective, A: Approach, R: Reference; P: Page in dissertation, (C) Conference article

C	Application	DBA Short Description	R	P
IV	Track and Trace	O: Traceability of material and products (e.g., position, cycle times).	12	76
		A: Track and trace of containers, materials, and products by using unique identifiers such as RFID tags.	13	
	Material Flow Management	O: Demand-oriented, automated control of the material flow.	14	77
		A: Two variations: (1) Push approach: a central production control system knows the demand, the position, and the availability of materials. An algorithm calculates the optimal material flow to meet demands with little material handling effort and low WIP inventories.		
		(2) Pull approach: a digital Kanban system (e-Kanban) detects automatically if the available material falls below a minimum threshold and triggers the replenishment process.	15	
	Inventory Management	O: Smart inventory management to ensure material availability with minimal inventory. A: Accurate tracking of inventory enables low stocks and timely re-ordering from the supplier.	16	78
V	Product Quality Monitoring	O: Identification of defective products (reactive quality control). A: Comparison of real-time quality data (e.g., geometrical dimensions) with reference values.	17	79
	Product Quality Improvement	O: Systematic and preventive avoidance of errors. A: The availability of accurate quality-related data enables failure root cause methods to identify and eliminate systematic error causes.	18	78
VI	Energy Monitoring	O: Reduction of energy consumption. A: Measuring energy consumption to find saving potentials.	19	82
	Environmental Monitoring	O: Ensure healthy working conditions for employees. A: Measuring environment conditions such as air quality.	20[1]	83
References		12. Louw and Walker (2018) (C), 13. Segura Velandia, Kaur, Whittow, Conway, and West (2016), 14. Jarupathirun, Ciganek, Chotiwankaewmanee, and Kerdpitak (2009), 15. Wan et al. (2018), 16. Saygin and Sarangapani (2006), 17. Wuest et al. (2014), 18. Oliff and Liu (2017), 19. Lenz, Kotschenreuther, and Westkaemper (2017), 20. Pilloni (2018)		

Cat: Category, O: Objective, A: Approach, R: Reference; P: Page in dissertation, (C) Conference article

[1] DBA is described as part of the article among other applications, no focused case study.

4.1.4 Data-based Applications Description

This chapter provides more information on each DBA shown in Figure 15. The reference in column *P* of Table 15 refers to the page in this work. The chapters are structured as follows: first, a brief description of the category is given. Second, each DBA of this category is presented, including a general description, a use case example, as well as its objectives and key requirements.[18] As stated before, the DBA notation is derived by the author following the primary purpose of the DBA. Hence, the notion used in the dissertation might differ from the referenced literature, however, it describes the same concept.

4.1.4.1 Planning & Scheduling

The first category is concerned with production preparation. It comprises two DBAs: *Layout Planning* and *Production Scheduling*. *Layout Planning* is infrequently used but might have a significant impact as it defines the production layout and thus imposes physical constraints. *Production Scheduling* is used iteratively to define the production sequence.

4.1.4.1.1 PRODUCTION SCHEDULING

General Description

Production Scheduling determines the process of orders and assigns key production resources, such as machines, to the production orders. The objective is to derive an optimal production plan, for example, maximizing asset utilization or minimizing cycle time while meeting the constraints of the manufacturing process. Therefore, data-based *Production Scheduling* takes manufacturing resource data—such as available equipment and capacity data, as well as material data—and other constraints into consideration (Zhu, Qiao, & Cao, 2017).

Use Case Example

Vallhagen et al. (2017) present a *Production Scheduling* use case for bottleneck resources in an aircraft component manufacturer. Part of the production process of several products is the heat treatment, which is a shared resource for internal and external products. It is highly utilized, and long waiting queuing times have occurred frequently. Therefore, reducing waiting times by increasing asset utilization has been a key objective. The heat treatment facility comprises six furnaces with different characteristics. The process of scheduling the furnaces is very complex, and efficient scheduling was difficult. In the past, the heat treatment process has been characterized by frequent priority changes, rescheduling, and waiting times. In order

[18] Key requirements are general and high-level requirements, such as machine learning skills, without going into technological details.

to determine an optimal scheduling plan while considering all constraints, the company introduced a mathematical optimization solver into the scheduling process. The optimization model integrates various kinds of data from the ERP system, manufacturing control system, and manual input. The scheduling optimization model has substantially contributed to a higher utilization rate of the furnaces and thus reduced queuing times.

Main Objectives

The main objectives of this application are an accurate and fast planning process, the ability to reschedule in case of disturbances, and high asset utilization.

By integrating several types of information, plans derived by *Production Scheduling* systems are more accurate than human-made plans (Qi & Tao, 2018, p. 3587). It also improves the planning speed and reduces human effort (Ray Y Zhong, Xun Xu, & Lihui Wang, 2017, p. 12). Zhong et al. (2014, p. 827) emphasize the significance of near real-time and adaptive production scheduling as in real-life production environments, unexpected disturbances may occur and make fixed plans obsolete. The ability to increase asset utilization is demonstrated in the presented use case.

Key Requirements

To derive accurate plans, *Production Scheduling* requires accurate and up-to-date data of the manufacturing system. For instance, planning tools often rely on standard operation times (SOTs) that are calculated based on experience. However, due to variations in the production process, these SOTs do inadequately reflect the real operating times. Zhong et al. (2014, p. 827), therefore, suggest applying RFID technology for real-time tracking of material and products to provide more realistic SOTs. This increases the accuracy of the scheduling process and allows for quick adjustments in case of disturbances. This view is supported by Landherr, Schneider, and Bauernhansl (2016, p. 30), who argue that having online data on production processes and logistic components available allows a faster reaction to unexpected events. To enable frequent updates of data, Vallhagen et al. (2017, p. 669) propose to automate data collection as much as possible.

A second key requirement is an in-depth process understanding. *Production Scheduling* often applies a mathematical optimization model. As each production area is different and has different constraints, a distinct optimization model needs to be designed for each area (Vallhagen et al., 2017, p. 669). The ability of the optimization model to adequately represent the shop floor reality thereby depends on the domain knowledge of the involved team.

4.1.4.1.2 LAYOUT PLANNING

General Description

The physical layout of the plant is a critical part of efficient and effective production (Kumar et al., 2018, p. 644). Production lines and material supply should be designed to achieve the highest level of productivity, safety, and quality (Kadane & Bhatwadekar, 2011, p. 59). Data-based layout planning can support the derivation of a layout plan that allows low WIP inventory, minimal material handling effort, and high asset utilization while meeting all layout constraints such as available space, emergency path space, or maximum floor weight. The optimization criteria of the model can be weighted to derive layout plans that meet the specific needs of the organization (Kumar et al., 2018).

Use Case Example

Kumar et al. (2018) present a use case of data-based layout design planning with a focus on sustainability. The layout planning design and selection comprises three phases. In the first phase, 14 layout criteria have been identified from the literature and expert discussion. With a principal component analysis, the 14 criteria were reduced to four factors: material handling distance, maintenance, adjacency, and hazard. To ensure environmental sustainability, the authors added electric energy consumption as the fifth factor in the optimization model. In the second phase, input data concerning the number of machines, products, and production cells is computed as basis for the optimization model. Variation of the production process in terms of product volume and used production equipment was modeled by computing 10 different production settings. Afterward, a mathematical optimization model is formulated and based on the input data, 10 different layout plans generated. The third phase comprises the evaluation and ranking of the layout. The authors applied three different techniques to rank the proposed layouts, resulting in similar but not equal rankings. Therefore, the final selection was still to be done by humans.

Main Objectives

Dependent on targets, different optimization criteria are selected. Therefore, the application provides a range of possible objectives, most notably WIP inventory reduction, movement reduction, high utilization of space and equipment, and short distances (Kumar et al., 2018, p. 644).

Key Requirements

To derive a layout plan that matches the real-world conditions, the optimization model requires the input of manufacturing data. Among these are volume data, such as the number of products and machines, variety data, including the number of different product types and data related to demand as well as SOTs (Kumar et al., 2018, p. 645).

4.1.4.2 Production Control

The general purpose of the category *Production Control* is to provide transparency on the performance of the manufacturing system to supervisors. Thereby, the two perspectives of the DBA *Real-time Control* and the DBA *System Performance Measurement* are differentiated.

4.1.4.2.1 REAL-TIME CONTROL

General Description

The basic idea of *Real-time Control* is to get instant feedback in case of deviations of the manufacturing process. The real-time comparison between sensor data from the physical production system and reference data allows ubiquitous process control, such as being able to detect deviation with very low latency times (J. Manyika et al., 2011, p. 81; Markarian, 2018, p. 11). Data-based production control is more and more relevant as production processes become increasingly complex and thus problems increasingly difficult to detect. Qi and Tao (2018, p. 3591) argue that “these visible and invisible problems in smart manufacturing can be reflected by the data.” According to James Manyika et al. (2015, p. 4), anomaly detection and real-time control are some of the most applied forms of data utilization on the shop floor.

Use Case Example

Hirmer et al. (2017) present a use case of *Real-time Control* in the context of computer server surveillance. The basic approach, called SitOPT,[19] however, is also applicable to manufacturing settings. SitOPT proposes an architecture system consisting of three layers: *Sensing* (1)*, Situation Recognition* (2)*,* and *Situation-Aware Workflow* (3). The Sensing layer comprises physical sensors, monitoring the environment and providing data for the upper layers. The Sensing Layer aggregates and processes data from layer 1. Anomalies in the process (called *situations*) are assessed based on so-called *Situation Templates.* Based on the situation identified in layer 2, the *Situation-Aware Workflow layer* either sends a notification to the system administrator or issues a predefined action to mitigate the situation.

Main Objectives

The main objective of this application is ubiquitous process control to detect deviations (James Manyika et al., 2015, p. 4) and to provide immediate feedback to the operator in case of deviations in the production process or to initiate a prefined mitigation action (Hirmer et al., 2017, p. 176).

[19] SitOPT is the name of a research project of the German Research Foundation (DFG) at the university of Stuttgart.

Key Requirements

Yan et al. (2017, p. 23484) emphasize the relevance of using advanced sensors and IT systems to ensure high availability of up-to-date and high-quality data from the shop floor. Pilloni (2018, p. 3) adds that real-time control of industrial processes requires stable data connections with low latency. Thoben et al. (2017, p. 12) point out that collecting and transferring a wide range of manufacturing data increases the motivation for external parties to gain access to the IT system illegally and therefore adds high IT security standards as another key requirement.

4.1.4.2.2 SYSTEM PERFORMANCE MEASUREMENT

General Description

In contrast to *Real-time Control*, the application *System Performance Measurement* has a more long-term and holistic perspective. The DBA provides a data-based overview of the current overall performance of the manufacturing system and allows trends to be identified. Therefore, key performance indicators (KPIs) are calculated automatically and visualized on a dashboard (Illa & Padhi, 2018, p. 55165). Automatically generated KPIs are an important element of performance measurement in the context of digital shop floor management (Meissner et al., 2018, p. 83).

Use Case Example

Friedemann et al. (2016) present a use case of productivity indicator calculation. One of the most established KPIs in production is the Overall Equipment Efficiency (OEE) rate. The OEE is the de facto standard for measuring equipment availability. Based on track and trace data, a so-called Cycle Time Analysis (CTA) is performed. This KPI not only serves as an indicator for cycle time but also as input for an automated OEE calculation. In the context of automated KPI calculation, Friedemann et al. outline the scenario of employees equipped with wearables displaying the KPIs of the machine the employee is looking at.

Main Objectives

The application has two main objectives. First, to provide a holistic overview of the current overall performance of the system and second, to support the production supervisor in detecting trends within performance metrics. This increases transparency and allows earlier detection of problems in the process. Automated calculation of KPIs reduces the effort and enables regular benchmarks within the plan or the production network (Meissner et al., 2018, p. 84).

Key Requirements

The calculation of KPIs requires accurate data from several machines in the production process and therefore has similar requirements as the application *Real-time Control* regarding sensors, data connection, and data security.

Meissner et al. (2018, pp. 83–84) argue that besides the technology basis, two more requirements must be met. First, while new technology minimizes the effort to calculate KPIs, management must resist the temptation to track too many KPIs. Second, the automation of data collection and processing bears the risk that shop floor workers feel disconnected from the performance measurement process and therefore leads to alienation from shop floor management. As a consequence, it is crucial for employees to understand how automatically generated KPIs are calculated and how they are to be interpreted.

4.1.4.3 Maintenance

Maintenance increases process stability by ensuring equipment availability (Lodewijks, Kraijema, Godjevac, & Corman). Insufficient or late maintenance has negative impacts on process performance and product quality (Sipsas, Alexopoulos, Xanthakis, & Chryssolouris, 2016, p. 236). Maintenance is critical for equipment uptime but is itself a significant cost driver, and may account for more than 50 percent of the machine lifecycle cost (O’Donovan et al., 2015b, p. 2). Therefore, maintenance staff is challenged by the trade-off between increasing equipment availability and minimizing maintenance expenditures (Z. Li et al., 2017, p. 377).

The importance of maintenance continues to grow because of the forecasted emergence of smart factories. The technical complexity of SM requires a comprehensive maintenance program to avoid unplanned breakdowns (Mayr et al., 2018, p. 625). O’Donovan et al. (2015b, pp. 2–3) add that the customer-focused and highly optimized supply chain of SM increases the need for high equipment availability.

The literature differentiates four kinds of maintenance policies: corrective maintenance, preventive maintenance, predictive maintenance, and prescriptive maintenance. In addition, the literature presents condition-based maintenance. Corrective maintenance, also called run-to-failure maintenance, is the simplest maintenance policy. As long as the equipment does not show a defect, it is assumed that it is in good condition. Corrective maintenance is purely reactive and does not follow any maintenance plan. Preventive maintenance is currently the most popular maintenance policy. It is a periodic policy that follows a maintenance plan that was designed at the moment of the equipment installation. The basic assumption is that the equipment will break if not maintained as planned (Munirathinam & Ramadoss, 2014, p. 895). Fixed intervals, however, come with disadvantages. Intervals that are too long bear the risk of unexpected breakdowns, and intervals that are too short cause unnecessary maintenance costs (Liao & Wang, 2013, p. 225).

The maintenance policies of condition monitoring, predictive, and prescriptive maintenance include the collection and analysis of data and are therefore considered DBAs.

4.1.4.3.1 CONDITION MONITORING

General Description

Condition Monitoring allows the monitoring of load, wear, and defects during operations. Due to continuously monitoring the current health situation of the production equipment, defects are detected earlier, resulting in less machine downtime and the prevention of consequential damages (Mayr et al., 2018, p. 625). *Condition Monitoring* may include, among others, the monitoring of noise and vibration. Anomalies in those features can be detected and used to identify equipment trouble or failure and thus, make better decisions in terms of maintenance and parts replacement (X. Li et al., 2017, p. 34). *Condition Monitoring* helps to overcome the disadvantages of fixed intervals, namely the risk of unexpected breakdowns and the risk of investing unnecessary maintenance costs (Liao & Wang, 2013, p. 225). Thereby, *Condition Monitoring* not only increases the effectiveness, but also the efficiency of maintenance activities.

Use Case Example

Yunusa-kaltungo and Sinha (2017) present a use case of *Condition Monitoring* on industrial rotating machines. They developed a vibration-based method to classify equipment faults. In the first step, vibration data was monitored from two machines, at three different speed levels and under six different states of health condition for 20 weeks. By clustering the vibration data, six distinct patterns of vibration are identified. These patterns are used to classify new vibration data, thus allowing faster identification of abnormal behavior. In this use case, the authors focused on vibration data only but highlight the potential to gain machine health information by monitoring other data, such as temperature, power consumption, sound, or the condition of lubricants.

Main Objectives

Summarizing the general description and the use case example, *Condition Monitoring* has at least three objectives. First, to increase equipment uptime and reliability by reducing unexpected breakdowns. Second, to increase process stability due to higher equipment reliability. Third, to reduce costs and minimize resource-intensive excessive maintenance by initiating maintenance activities only if condition monitoring data suggests an abnormal behavior of the equipment.

Key Requirements

The use case suggests at least three key requirements. First, sensors at the production equipment are necessary to collect equipment health data, such as vibration, temperature, or power consumption data. New machines usually already integrate a range of sensors. However, condition monitoring is not limited to brand-new equipment, but existing equipment can be upgraded with sensors at reasonable costs (Schöning & Dorchain, 2014, p. 545). Second, as the monitoring data needs to be stored and processed, an appropriate IT infrastructure is required. And third, to facilitate meaningful interpretations of visualizations (e.g., vibration clusters), employees need both a general process understanding as well as technical expertise of the monitored equipment.

4.1.4.3.2 PREDICTIVE MAINTENANCE

General Description

Predictive Maintenance is going beyond the concept of monitoring equipment conditions to find anomalies. Based on historical and real-time data, indicators of anomalies are predicted even before they happen. By performing maintenance activities based on the forecasts preventively, unexpected breakdowns are reduced and equipment life is maximized (Munirathinam & Ramadoss, 2014, p. 896). Liao and Wang (2013, p. 225) have observed a shift of many manufacturers from traditional corrective and preventive maintenance to predictive maintenance.

Use Case Example

Z. Li et al. (2017) present a *Predictive Maintenance* use case for mechanical systems based on DM in a smart factory environment. They describe a system framework for fault diagnosis and prognosis in machine centers. The framework consists of the following five modules: The sensor selection and data acquisition module (1) includes the selection and installation of suitable sensors as well as the definition of the data collection strategy. As the collected data contains much irrelevant information, it enters the data preprocessing module (2), including data cleaning, data integration, data reduction, and data transformation. The DM module (3) is responsible for failure detection, classification, and failure forecasting. In the use case, the backlash error between screw and table is the indicator of the need for imminent maintenance. The research team collected data on the backlash error for 22 weeks without maintenance. Seventy percent of the data was used to train an artificial neural network, the rest to test the artificial neural network and to forecast the backlash error. The decision support module (4) visualizes the results. Finally, the maintenance implementation module (5) derives a maintenance plan.

Main Objectives

Predictive Maintenance aims to find the optimal balance between too much and too little maintenance, thus ensuring a required level of equipment uptime while minimizing the necessary maintenance costs (Z. Li et al., 2017, pp. 377–378). Going beyond the concept of condition monitoring, *Predictive Maintenance* intends to predict the future conditions of the equipment. By forecasting the remaining useful life of a machine element, the application can determine the best time for maintenance activities more accurately than fixed schedule maintenance approaches.

Key Requirements

Predictive Maintenance depends on accurate and up-to-date machine condition data. By monitoring and analyzing a variety of data over an extended period, predictions about the failure probability of critical components are derived. A prescriptive indicator for the need for maintenance is, for instance, slow changes in vibration intensity (Schöning & Dorchain, 2014, p. 545). Several authors explicitly highlight the essential relevance of manufacturing expert knowledge for *Predictive Maintenance*. Data analysts struggle to select the correct data sources and to derive well-founded hypotheses or optimization models without having access to the sound process understanding of domain experts (Åkerman et al., 2018, p. 416; Lee, Kao, & Yang, 2014, p. 7; Liao & Wang, 2013, p. 229; Mayr et al., 2018, p. 625).

4.1.4.3.3 PRESCRIPTIVE MAINTENANCE

General Description

Prescriptive Maintenance is an extension of the concept of *Predictive Maintenance*. Following the concept of the four data analytics capabilities of Gartner (see chapter 2.3.4), the first three levels *descriptive*, *diagnostic*, and *predictive* merely provide the foundation for informed decisions of humans. The prescriptive analytics level, in contrast, already proposes, or even initiates, an action. Transferring the concept of prescriptive analytics to maintenance, *Prescriptive Maintenance* aims to derive data-based decisions about the optimal maintenance plan independently from humans (Bokrantz, Skoogh, Berlin, & Stahre, 2017, p. 165).

Use Case Example

As the overview of DBA in key references (see Table 14) has shown, *Prescriptive Maintenance* is currently no relevant part of the maintenance discussion in the literature. Only one paper addressing the application was found in the literature review. Matyas et al. (2017) present a use case of *Prescriptive Maintenance* applied to a triaxial machining center in an automotive facility. Based on incoming real-time machine data and identified correlations, a mathematical reaction model predicts machine conditions. A software tool called “Prescriptive maintenance decision support system” (Matyas et al., 2017, p. 463) visualizes the predicted trends and provides suggestions for anticipative maintenance measures. The authors report that the system was able to predict 43 percent of unplanned machine breakdowns caused by mechanical failures, thus leading to substantial cost savings. On the other hand, the design of the 3D model of the machine, serving as the basis for the reaction model, is very resource consuming. While the authors do not specify the logic behind the suggested maintenance measures, they point out that the system cannot substitute the knowledge of operators.

This system matches the first of the two levels of prescriptive analytics (see Figure 5). A use case of *Prescriptive Maintenance*, including decision automation, was not found in the literature. As of today, AI is not able to substitute domain expertise adequately to justify a shift from human-based decision-making to automated decision-making.

Main Objectives

The DBA *Prescriptive Maintenance* has the same main objectives as the DBA *Predictive Maintenance* in terms of equipment reliability, process stability, and maintenance efficiency. The DBA *Prescriptive Maintenance* provides data-based support to operators by suggesting anticipative maintenance measures. The final stage of *Prescriptive Maintenance* will not only derive maintenance decisions but also initiates activities without human confirmation.

Key Requirements

Very similar to the key requirements of *Prescriptive Maintenance*, the two most essential requirements for *Predictive Maintenance* are access to relevant data and employee know-how. This includes manufacturing domain knowledge and data science skills (Nemeth, Ansari, Sihn, Haslhofer, & Schindler, 2018, p. 1039), including exploratory data analysis methods and ML (Nemeth et al., 2018, p. 1040).

4.1.4.4 Internal Logistics

This category is concerned with the tracking, transport, and storage of raw materials, intermediates, and final products. It comprises three DBAs with a distinct scope. The DBA *Track and Trace* collects a variety of product-related data and is thus the precondition for more sophisticated applications. The second DBA *Material Flow Management* seeks to optimize the flow of material within the factory, and the third DBA *Inventory Management* aims to derive an optimal level of warehouse inventory.

4.1.4.4.1 TRACK AND TRACE

General Description

The purpose of the DBA *Track and Trace* is described by its name. Unique identifiers such as RFID tags are used to track and trace the real-time status (e.g., position within the plant) of various objects in manufacturing, such as material, container, and products (Ray Y Zhong et al., 2017, p. 2). Tracking data is collected in real-time and uploaded in the central IT system, thus increasing the transparency of the internal material and product flow (Mayr et al., 2018, p. 624).

Use Case Example

Segura Velandia et al. (2016) present a use case of *Track and Trace* in crankshaft manufacturing. They proposed to use RFID tags to trace the product's life cycle and provide full transparency over the components process history. Full product history transparency reduces costly downtime to rectify processing defects and product recalls.

In the use case, the crankshafts are transported by a gantry system. The gantry arm is equipped with an integrated antenna reader, including an RIFD transmitter and receiver. The RFID tag is attached to the crankshafts directly. As the gantry arm is moving with the crankshafts through all process steps, one integrated antenna reader is sufficient to track the complete process history. The maximal distance of typical RFID tags and RFID reader are approximately six meters. The installation of the RFID system costs around 17,500 EUR, including 1,000 RFID tags, demonstrating that RFID technology is a cost-effective opportunity for track and trace.

Main Objectives

The main objective is to identify each object and to track its way through the production process. Thereby, the production history of each product can be traced back in case of quality issues, a feature that is often requested by OEMs from its suppliers (Segura Velandia et al., 2016, p. 67). *Track and Trace* allows the tracking of WIP, increases the transparency of the material flow, and might contribute to a reduction of safety stock (Mayr et al., 2018, p. 624).

Key Requirements

Track and Trace is enabled by RFID technology. RFID systems comprise three components: the RFID tag or transponder; the RFID reader; and a small computer unit that processes, stores, and transmits the data. As a comparably low effort implementation of an RFID system is feasible, it is also interesting for SME companies (Louw & Walker, 2018, p. 256). A challenge highlighted by Segura Velandia et al. (2016, p. 76) is the integration of RFID data into MES or ERP systems.

4.1.4.4.2 MATERIAL FLOW MANAGEMENT

General Description

The application *Material Flow Management* aims to provide input material to production equipment at the right time while minimizing material handling effort and WIP inventory. Two opposing strategies for material flow control can be distinguished—push and pull. A push system is the result of central production and material flow planning that follows deterministic rules. The use of a push system was driven by the emergence and nowadays widespread application of MES systems (Gerberich, 2011, p. 234). Pull is a demand-oriented material flow principle in which the material is only supplied if requested by the downstream process step (Womack & Jones, 2003).

Use Case Example

Kolberg and Zühlke (2015, p. 1871) present two use cases for IT technology enhanced material flow systems. The Würth Industrie Services GmbH & Co. KG introduced in 2013 an optical pull-oriented order system called iBin. The filling status of the bin is monitored by a camera system. The current level is reported wirelessly to an inventory control system, thus allowing a real-time assessment of the current stock of material stored in iBins. The system can be configured to automatically send orders for replenishment if a minimal level of inventory is reached.

The second use case presents a combination of pull and push. The Wittenstein AG and the Bremer Institut für Produktion und Logistik GmbH work on an IT-supported material supply system. Based on real-time demand a central IT system calculates intervals for milk-run supply rounds.

A third use case is described by Wan et al. (2018, p. 55421). Manufacturing data and AI enable intelligent path planning for automated guided vehicles (AGV), thus reducing transport times and energy consumption.

Main Objectives

The main objectives of this application are the timely delivery of required material at the production equipment, minimum WIP inventory, and safety stocks levels (Mayr et al., 2018, p. 624) as well as minimal handling and transportation effort due to intelligent path planning (Wan et al., 2018, p. 55421).

Key Requirements

Both presented approaches—push and pull—depend on highly accurate data of current stock availability and demand. New technologies, such as smart sensors and RFID tags, collect accurate data of the current material flow. The push system benefits from the precise planning of the underlying planning tool due to the availability of high-quality planning data. The pull system benefits from track and trace and other material-related data, as transparency on material availability and current consumption allow for the optimal time for replenishment to be determined.

4.1.4.4.3 INVENTORY MANAGEMENT

Basic Description

The purpose of the DBA *Inventory Management* is to balance the need for high availability of input material or finished products on one hand, and low inventory levels to minimize stock keeping costs on the other hand (J. Li et al., 2015, p. 677).

Use Case Example

Saygin and Sarangapani (2006) provide a use case of RFID optimized inventory management of time-sensitive materials in a manufacturing company. They compare the performance of three inventory management models in terms of service level and expired materials. The first two models are static and rely on fixed baseline inventory levels for replenishment. The third model integrates a dynamic inventory forecast based on track and trace data. It considers the difference between the current inventory level of an object at a storage area and the predicted demand to determine the right amount of material for replenishment. The results suggest that the RFID data-based dynamic model can adapt more effectively to system dynamics than fixed baseline-oriented inventory management approaches while holding reduced levels of inventory.

Main Objectives

The main objective of *Inventory Management* is to increase the transparency of the current inventory status (Sanders et al., 2016, p. 823) to “replace inventory with perfect information” (Clegg & Powell, 2013, p. 1497). Due to full transparency of inventory and demands, the objectives of material availability and minimal stocks can be balanced more effectively.

Key Requirements

The key requirement of *Inventory Management* is full transparency on current inventory, planned production, and expected replenishment. This translates to the requirement to accurately measure the incoming and outgoing material to assess the current inventory status (J. Li et al., 2015, p. 677). This requirement can be met with RFID technology. To allow inventory forecasting, current inventory data needs to be combined with data on replenishment orders and production planning data. Kletti (2015, p. 98) adds that to have real-time transparency on the inventory, tracking data needs to be fed back into the MES or ERP system with low latency.

4.1.4.5 Product Quality Management

The category *Product Quality Management* seeks to provide customers with flawless, high-quality products. This category comprises two DBAs: the DBA *Product Quality Monitoring* monitors product quality and detects defective products, and the DBA *Product Quality Improvement* uses quality-related data to identify root causes of defects.

4.1.4.5.1 PRODUCT QUALITY MONITORING

Basic Description

Product Quality Monitoring aims to identify defective parts or finished products and thus prevent defects from being passed on to the next process step or even to the customer. By monitoring product quality, deviations from the nominal quality characteristics can be identified and either corrected automatically or via manual intervention (Majstorovic et al., 2018, p. 504). For example, machine vision, including visual inspection per camera system and image processing algorithms, is used to detect scratches on the product's surface. Machine vision is a high-reliability, efficient, and accurate application to monitor visible product quality characteristics (Wan et al., 2018, p. 55421).

Use Case Example

Wuest et al. (2014) present an approach to quality monitoring in manufacturing using supervised machine learning. First, cluster analysis is used to define a finite set of desired product states. Afterward, ML is used to monitor several product quality characteristics simultaneously to detect product states that are not within the defined set of desired states. The objective of the approach is to trigger an action if a product state shows too much variation from the desired state. This might include notifying a quality inspector to adjust the manufacturing process or initiate the scrapping of the product to avoid passing on the defective part to the next process step. Another example of product quality monitoring is surface inspections, to test whether geometric parameters (e.g., surface roughness) are in accordance with specifications (Tao et al., 2018, p. 9).

Main Objectives

The main objectives are threefold: the first objective is to ensure the production and delivery of defect-free products to the customer. Furthermore, by avoiding defective products being passed on to the next process step, the quality costs for late defect detection are reduced. Thirdly, data of monitoring product quality form the basis for consistent reporting on quality issues (Gewohn et al., 2018, p. 459). Quality data can be used to identify and rank failure clusters and thus support quality inspection as a basis for a data-based root cause analysis.

Key Requirements

Based on the use case example, three requirements can be derived for quality monitoring. First, sensors and cameras are required to monitor product quality characteristics. Monitored data needs to be compared to a set of reference values. Therefore, a database comprising thresholds of reference values as well as reoccurring defects needs to be established. To ensure an accurate and up-to-date quality database, Gewohn et al. (2018, p. 459) suggest consistent quality feedback processes in real-time. Finally, depending on the complexity of the quality monitoring approach, various skill sets of employees are required. The presented ML approach needs expertise in statistics (for cluster analysis) and ML programming as well as manufacturing domain expertise to interpret the results of statistical analysis.

4.1.4.5.2 *PRODUCT QUALITY IMPROVEMENT*

General Description

In contrast to *Product Quality Monitoring*, which is a reactive element of product quality management, the DBA *Product Quality Improvement* seeks to learn from failures in the past, to identify defect root causes systematically and derive measures to ensure quality preventively. As Harding et al. (2006, p. 973) argue, a common approach to solving quality problems is to examine past quality issues, to better understand the process, and use this knowledge to improve the system to minimize future quality problems. Knowledge gained from analytics of manufacturing data, especially product quality data, can be integrated with knowledge-based systems to support product quality improvement (Harding et al., 2006, p. 973).

Use Case Example

Oliff and Liu (2017) present a use case of data-based *Product Quality Improvement* at a company producing washing machines. The provided dataset on previous failures includes information on the fault group and details on the specific nature of the fault. Based on this information, a rule-based learning algorithm was used to derive rules for failure classification. After using a sample of 5,000 failures for training the algorithm, the rate of correct classification of new failures achieved almost 95 percent. The failure classification is used to improve the design of products and make future quality monitoring more efficient.

Lokrantz, Gustavsson, and Jirstrand (2018) provide a conceptual paper on ML techniques for root cause analysis of quality deviations in manufacturing. Bayesian Networks are selected to construct models that draw conclusions on the root cause of quality deviations. Besides explaining the concept of Bayesian networks as a technique to identify causes for deviation, the authors demonstrate the importance of expert knowledge of interdependencies of manufacturing process elements. Bayesian networks, including expert input, perform much better in terms of failure detection accuracy than networks without expert input, regardless of the size of the training set.

Main Objectives

The main objective of *Product Quality Improvement* is to exploit quality-related data to understand product defects and derive systematic measures to prevent those defects from reoccurring (Illa & Padhi, 2018, p. 55169). Failure detection and prevention leads to a higher quality of the final product and decreases costs for rework and scrapping of defective parts.

Key Requirements

Requirements for the DBA *Product Quality Improvement* are similar to those of the previous application. The collection and storage of quality-related data is the basis for subsequent analysis for failure clustering and root cause analysis (Ngo & Schmitt, 2016, p. 499). Depending on the actual application, different skills are necessary. For instance, DM skills are required for a classification system and image processing techniques know-how for surface inspections. To ensure a meaningful interpretation of quality data and to evaluate the outcome of data analytics, employees require process and product specific knowledge.

4.1.4.6 Environment, Health, and Safety

The category *Environment, Health, and Safety* has two primary purposes: to monitor and reduce energy consumption, and to ensure healthy and safe working conditions by monitoring the working environment.

4.1.4.6.1 ENERGY MONITORING

General Description

Energy costs constitute a significant part of the overall costs of manufacturing. Almost all production equipment—such as machines, robots, and sensors—consume energy. *Energy Monitoring* aims to measure energy consumption, to identify energy usage patterns, and to derive measures to reduce the overall energy consumption (Illa & Padhi, 2018, p. 55169).

Use Case Example

Lenz et al. (2017) present a use case of *Energy Monitoring* of manufacturing equipment. The objective is to reduce energy consumption by quantifying the current consumption of different machine components. The current energy consumption is visualized per period and per component. The visualization enabled the research team to identify components that unexpectedly consume too much energy. For instance, in the use case, the main consumer of energy are auxiliary units, such as the hydraulic system. Energy consumption reduction is achieved by switching to more efficient auxiliary units. Furthermore, the authors found considerable saving potential of energy by using the sleep mode of idle components more frequently. In total, the authors reported a decrease in energy consumption of 28 percent.

Main Objectives

The main objective is to detect sources of unnecessary energy consumption and to minimize overall energy consumption.

Key Requirements

According to Lenz et al. (2017, p. 366), the energy consumption of the machine is measured by analyzing the signals from a power monitor PLC (programmable logic controller). The accurateness thereby depends on the sampling rates of the PLC.

4.1.4.6.2 ENVIRONMENTAL MONITORING

Basic Description

The application *Environmental Monitoring* seeks to monitor the production environment to provide a safe and healthy working space for the shop floor employees (Pilloni, 2018, p. 6).

Use Case Example

Pilloni (2018) discusses workplace safety as one aspect of Big Data in industry 4.0. Accordingly, smart devices, such as helmets and watches equipped with sensors, are used to detect workplace hazards before causing any harm. For instance, sensors in helmets can monitor the concentration of gases, such as CO2, SO2, and SH4 in the air and raise an alarm if reference values are exceeded.

Main Objectives: The application's main objective is to ensure a safe and healthy working space.

Key Requirements: The prerequisite for the presented use case are sensors that monitor the environment and can be carried by humans without much effort (Pilloni, 2018).

4.1.5 Core Functions

4.1.5.1 Introduction and Overview

The previous chapter has indicated a broad spectrum of DBAs with different objectives (see Table 14). However, abstracted from the individual goals, all DBAs comprise one or more of four core functions, which are Monitoring (1), Deviation Control (2), Decision Support for humans (3), and Autonomous Optimization (4).

Table 16 depicts the core functions per DBA.

Table 16: Core functions of data-based applications

Data-based Application	Core Function			
	Monitoring	Deviation Control	Optimization	
			Decision Support	Autonomous by DBA
I. Planning and Scheduling				
Production Scheduling				•
Layout Planning			•	
II. Production Control				
Real-time Control	•	•		
System Performance Measurement	•		•	
III. Maintenance				
Condition Monitoring	•	•		
Predictive Maintenance	•		•	
Prescriptive Maintenance	•			•
IV. Internal Logistic				
Track and Trace	•		•	
Material Flow Management	•			•
Inventory Management	•			•
V. Product Quality Management				
Product Quality Monitoring	•	•		
Product Quality Improvement	•		•	
VI. Environment, Health, and Safety				
Energy Monitoring	•		•	
Environmental Monitoring	•	•		

The four core functions are derived based on a comparison of the functional range of the DBAs identified and discussed in chapter 4.1.4. When comparing the scope of the individual DBA, it is evident that although they have different objectives, the underlying functions required to deliver the objectives are quite similar.

For example, the two DBAs *Real-time Control* and *Condition Monitoring* have different objectives (detection of process anomalies vs. reduction of equipment breakdowns). However, on an abstract level, both applications collect data and compare the data to reference data of a desired state; thus, both share the functions (data) *Monitoring* and *Deviation Control*. Functions that have been identified recurrently in several DBAs are called core functions in this dissertation. An analysis of all 14 DBAs revealed that their functionalities can be described by the four core functions presented in Table 16.

For example, as seen above, *Condition Monitoring* can be described by the two core functions Monitoring and Deviation Control. *Predictive Maintenance* monitors data too but also applies prediction models to forecast the future condition of a machine. Based on the forecast, the maintenance staff derives optimized maintenance plans. Hence, *Predictive Maintenance* comprises the two core functions *Monitoring* and *Decision Support* for humans. In the vision of *Prescriptive Maintenance*, historical and current equipment health data is used to predict future equipment conditions and derive maintenance plans automatically. Therefore, *Prescriptive Maintenance* combines the two core functions *Monitoring* and *Autonomous Optimization*.

4.1.5.2 Characterization

This chapter provides a characterization of the four DBA core functions introduced above. [20]

1. Monitoring

The first core function is *Monitoring*. It includes the monitoring of manufacturing system elements, such as machine and material. Data is collected and stored and thereby provides the basis for more advanced functions. The primary objective of this function is to increase transparency by providing accurate and, if required, real-time data from the manufacturing system.

Monitoring is a core function of all the DBAs except for those of the category *Planning and Scheduling.* Monitoring is a basic function. The three functions Deviation Control, Decision Support, and Autonomous Optimization build on the data of the basic

[20] The concept of core functions in the context of data utilization in manufacturing and the distinction of the four core functions has been derived by the author of this dissertation based on the information presented in chapter 4.1.4. Consequently, no reference to external literature is given in the characterization of the four core functions.

function Monitoring to achieve a specific purpose and are therefore referred to as advanced core functions.

2. Deviation Control

The second core function is *Deviation Control*. Data from the basic function Monitoring is compared in real-time to reference values, and a notification is given if the monitored data are outside predefined thresholds.

Deviation Control is a core function of the following applications. *Real-time Control* reports deviation in the manufacturing process and highlights the position of the deviation. *Condition Monitoring* detects abnormal behavior of machines and informs maintenance personnel. *Product Quality Monitoring* identifies products which are not in specifications and sorts them out, and *Environmental Monitoring* raises the alarm if limit values for dangerous substances are exceeded.

3. Optimization – Decision Support

The third core function *Decision Support* enables manufacturing employees to make more informed decisions. It comprises the analysis and visualization of manufacturing data. Data analysis may include finding correlations and patterns and based on that, the evaluation of different options. For instance, the DBA *Layout Planning* provides different layout design suggestions along with a ranking based on several criteria. However, the decision on the final layout is made by humans. *System Performance Measurement* automatically calculates and visualizes KPIs and their development over time. Thereby, it supports manufacturing employees to keep track of the performance of the system and to identify negative trends.

Predictive Maintenance forecasts a machine future condition, thus enabling maintenance personnel to derive better maintenance plans. *Track and Trace* visualizes the current flow of material and products, thus supporting employees to detect bottlenecks. *Product Quality Improvement* comprises different approaches, such as ML, to support employees in identifying the root cause of quality problems systematically. Finally, *Energy Monitoring* documents current energy consumption of a machine, which serves as a basis for employees to decide on actions for energy consumption reduction.

4. Autonomous Optimization

The fourth core function allows decision-making without human interaction. Based on a mathematical prediction or optimization model, decisions aiming at optimizing the manufacturing system are derived automatically.

Production Scheduling automatically derives the production plan, based on customer orders and available resources. *Prescriptive Maintenance* seeks to identify patterns in machine data to forecast future machine conditions and, going beyond *Predictive Maintenance,* derive maintenance activities autonomously. *Material Flow Management* organizes an optimal material flow based on material demand and availability, and *Inventory Management* ensures an optimal level of inventory by triggering replenishment from suppliers automatically.

The advanced core functions build on the basic core function. Except for *Production Scheduling* and *Layout Planning*, which rely on central data from the ERP system, all DBAs require at least near real-time data from the manufacturing system. Therefore, almost all DBAs include *Monitoring* as a basic core function. *Monitoring* as a standalone function, however, does contribute little value to the manufacturing system. Only by utilizing the data for a specific purpose creates value that justifies data collection in the first place.

4.1.5.3 Summary

Generally speaking, manufacturing data can be used for three purposes (see Table 16). The function *Deviation Control* compares manufacturing data to threshold values. This function is of comparably low complexity and comparably easy to implement. However, *Deviation Control* does not contribute to continuously improving the manufacturing system. Instead, it supports maintaining the status quo. Hence, the potential added value of the function *Deviation Control* to the manufacturing system is limited.

The core function *Decision Support* seeks to utilize manufacturing data to find a better solution as the status quo. In contrast to the first two core functions, the third and fourth core functions *Optimization – Decision Support* and *Autonomous Optimization* include data analytics. Following the definition presented in chapter 1.3, data analytics is “a scientific process of logical-mathematical transformation of data to improve decision-making,” DBAs that include the third and fourth core functions, therefore, form the subgroup of *(data) analytics DBAs.*

For example, the DBA *Predictive Maintenance* uses data analytics techniques to identify patterns within equipment data. Thereby, the DBA provides decision support to employees to derive an improved maintenance plan. However, the higher potential value add comes at the cost of a higher complexity of data analytics.

The most advanced core function is *Autonomous Optimization.* The promise of this function is a self-optimization of the manufacturing system by optimization models (e.g., to optimize the material flow) or ML applications to derive optimal maintenance plans automatically. Reducing the time and reducing human effort for decision-making by self-optimization may sound very promising for manufacturing companies. However, this function is linked to the highest complexity level of all four core functions. Mathematical optimization models and self-learning ML techniques require a high level of manufacturing and DBA specific expertise. Moreover, ML requires large samples of high-quality data to learn from.

In summary, two conclusions can be made. First, *Monitoring* is a basic core function that is needed as a prerequisite for the three other core functions but provides little value as a standalone function. Second, the three functions *Deviation Control*, *Decision Support*, and *Autonomous Optimization* use the monitored manufacturing data for a specific purpose and therefore create value for the system. However, there is also a tendency of higher complexity and effort that comes with the higher value of the more advanced core functions.

4.1.6 Key Requirements

This chapter consolidates the key requirements from the DBA descriptions in chapter 4.1.4 into the three categories Technical infrastructure, Data availability, and Know-how (see Table 17).

4.1.6.1 Technical Infrastructure

The DBAs document several different requirements related to the technical infrastructure. Data needs to be collected in the right quality, posing high requirements to the availability, accurateness, and reliability of sensors. Despite the increasing number of sensors, some data points (e.g., failure codes) are still entered manually. This kind of data collection is time-consuming and prone to errors. As far as technological and economically feasible, data collection should be automated. These requirements are summarized as key requirement *Data collection*.

Data needs to be up to date. DBAs, such as *Real-time Control*, rely on real-time data from different sensors. Therefore, these applications require a stable data connection with low latency times. These requirements are summarized as key requirement *Data transfer.* To this end, Waibel, Steenkamp, Moloko, and Oosthuizen (2017, p. 736) consider the incomplete presence of broadband expansion and the absence of a sophisticated mobile data network, as a barrier for many SM applications as the transfer rate is too slow and the latency times too high.

As data is originating from different sources and in different formats, it must be preprocessed before being used for applications. This includes data integration and transformation in specific formats. The process of data preprocessing is especially challenging for unstructured data, which is present in the majority of manufacturing systems (Wan et al., 2018, p. 55427). These requirements are summarized as key requirement *Data preprocessing.*

Finally, as Thoben et al. (2017, p. 12) point out, the vast amount of manufacturing data collected by the DBAs is an attractive target of criminals. IT systems must meet the highest safety standards to avoid data theft and sabotage acts. For this reason, cybersecurity is considered a critical element of SM (Ghobakhloo, 2018, p. 921). A survey among 126 SMEs has found that SMEs currently lack confidence in IT and data security, a fact that might hamper the introduction of SM technologies in these companies (Sommer, 2015, p. 1515). Hence, *Data protection* is added as another key requirement of DBAs.

4.1.6.2 Data Availability

Chapter 4.1.4 reveals differences between the DBAs regarding the need for historical data, real-time data, and central data. Historical data thereby refers to data that has been collected by the DBA itself in the past. For instance, *Predictive Maintenance* needs historical machine failure data to identify patterns for predictions. To characterize *real-time* data, the definition of real-time follows the understanding of the Cambridge dictionary. Accordingly, the term *real-time* is "used to describe the way in which a computer system receives data and then communicates it or makes it available immediately" (Cambridge University Press, 2019c). Thus, real-time data is data that is available very shortly after it was monitored. Real-time data is required, for example, by the application *Real-time Control.* The third type of data required is data that is usually available in central ERP or MES systems, including manufacturing orders, production plans, and inventory data. Central data is required, for instance, for the application *Production Scheduling*.

4.1.6.3 Know-how

The third category of DBA key requirements is *Know-how.* Several authors (Åkerman et al., 2018, p. 416; Lee et al., 2014, p. 7; Liao & Wang, 2013, p. 229; Mayr et al., 2018, p. 625) highlight the decisive importance of process understanding for the success of data utilization in manufacturing. Manufacturing domain expertise, such as product and maintenance know-how, is essential for several DBAs for two reasons. First, the quality of any model representing the shop floor condition depends on the ability to accurately describe the interdependencies of the process and thus requires domain knowledge. Second, when it comes to the interpretation of manufacturing data as well as results of data analytics, manufacturing domain expertise is indispensable.

Depending on the complexity of a DBA, specialized application expertise is needed. For instance, the application *Predictive Maintenance* requires expertise in DM or ML to identify patterns as a basis for predictions. These skills are currently rare among shop floor employees.

As a stable and performant IT infrastructure is the foundation for data collection, transfer, storage, and processing as well as for data protection, IT know-how is required to install, update, and maintain this system and is therefore also a key requirement of many DBAs.

4.1.6.4 Overview and Summary

Table 17 presents a consolidated overview of key requirements per DBA. The key requirements discussed above thereby serve as a reference and each DBA was evaluated against these requirements. For example, the requirement *Data protection* was stated explicitly only once, but in fact, applies to every DBA. Consequently, Table 17 exhibits this requirement as a requirement for all applications.

Advanced requirements refer to emerging SM technologies such as Cloud Computing or the next-generation mobile communication network 5G and are indicated by the symbol (●). Basis requirements refer to technologies that are today's standard in manufacturing and are indicated by the symbol (○). If a requirement is not applicable for a DBA, no symbol is used. The same concept is applied to the know-how category, where (●) indicates the need for advanced know-how and (○) the need for basic know-how. Taking a look at Table 17 allows one to draw three conclusions:

First, the majority of the requirements discussed in the use cases originating from literature focus on technological capabilities. While almost all articles discuss requirements of data collection, transfer, and processing, only a few include the aspect of the human factor as a critical enabler of DBAs. Only by highlighting the importance of manufacturing domain knowledge, authors indirectly discuss the role of the existing workforce for the implementation of DBAs. Similarly, organizational capabilities are scarcely present in the reviewed papers. This fact might be explained by the technological scope of the reviewed articles. Qualitative case studies in chapter 5, focusing on employee and organizational capabilities, will compensate for this shortage.

Second, in line with the observation of different complexities of DBA core functions, Table 17 indicates different levels of complexity of DBAs, reflected by different requirements. The DBAs *Track and Trace* and *Condition Monitoring* mainly require the existence of an appropriate technical infrastructure. Installing the sensors and tracking the respective data requires a financial investment but comparably low levels of manufacturing domain expertise and DBA specific expertise. As sensors can be bought and their installation be supported by suppliers, the basic requirements can be met in a relatively short period of time. More sophisticated DBA, such as *Material Flow Maintenance* and *Predictive Maintenance*, however, strongly depends on manufacturing domain knowledge and DBA specific expert knowledge. Therefore, those capabilities cannot be sourced externally but must be built internally. Building these capabilities requiring an existing pool of employees with sufficient prior knowledge and skills, as well as a comparable long time horizon, thus making the implementation of these DBAs more challenging.

Table 17: Key requirements originating from DBA use cases in literature

DBA Key Requirements											
Requirement Category		**Technical infrastructure**				**Data availability**			**Know-how**		
Requirement		Data collection	Data transfer	Data preprocessing	Data protection	Historical data	Real-time data	Central data	IT Know-how	Manufacturing domain expertise	DBA specific expertise
C	**Data-based Application**										
I	Production Scheduling			●	●		●	●	○	●	●
	Layout Planning			○	●			●		●	●
II	Real-time Control	●	●	○	●		●	●	○	○	
	System Performance Measurement	●	○	○	●	●		●	○	●	
III	Condition Monitoring	●	○	○	●		●		○	○	○
	Predictive Maintenance	●	○	●	●	●	●		○	●	●
	Prescriptive Maintenance	●	○	●	●	●	●		○	●	●
IV	Track and Trace	●	●	○	●				○	○	○
	Material Flow Management	●	●	●	●		●	●	○	●	●
	Inventory Management	●	●	●	●	●	●	●	○	●	●
V	Product Quality Monitoring	●	○	○	●	●		●	○	○	○
	Product Quality Improvement	●	○	○	●	●		●	○	●	●
VI	Energy Monitoring	●	○	○	●				○	●	○
	Environmental Monitoring	●	○	○	●				○	○	○

C: Category, ○ Basic requirements, ● Advanced requirements

Third, supporting the differentiation in basic and advanced DBA core functions in the last chapter, the more complex DBAs can build on the data foundation created by the rather basic DBAs. This observation can be made in four of six categories. The applications *Real-Time Control, Condition Monitoring, Track and Trace*, and *Product Quality Monitoring* may serve as data provider for the applications *System Performance Measurement, Predictive Maintenance, Material Flow Maintenance,* and *Product Quality Improvement* respectively.

4.1.7 Summary DBAs in Manufacturing

Data utilization is discussed in several distinct literature streams. As shown in Table 14, the literature covered in this chapter to identify DBAs includes literature dealing with ML, DM, big data in general, and publications related to SM and Industry 4.0. Likewise, the review has revealed a wide range of applications for DBAs. As depicted in Figure 15, six DBA categories consisting of 14 individual DBAs have been identified. These DBAs address several functions of a manufacturing system including production planning, production control, maintenance, logistics, and quality assurance. In addition, monitoring the environment can contribute to healthy working space and reduced energy consumption.

Based on the review of twelve identified articles on data utilization in manufacturing, Table 14 presents an overview of how often each DBA was mentioned in the articles. The following five DBAs have been discussed most frequently. *Production Scheduling* (discussed in 10 out of 12 articles), *Condition Monitoring* (9), *Predictive Maintenance* (10), *Product Quality Monitoring* (9), and *Product Quality Improvement* (9).

Since DBAs are applied in several functions of a manufacturing system, the individual objectives are quite diverse. Table 15 shows the individual objectives of each DBA, along with a short description of a possible approach to achieve the objective. However, abstracted from the individual goals, all DBAs are comprised of one or more four core functions. These are Monitoring, Deviation Control, Decision Support, and Autonomous Optimization. Monitoring is considered as a basic core function, as it is the foundation for the other three functions, but does not provide much value to the company as a standalone function. Value is only created by actually using the collected manufacturing data to serve a specific purpose. The core function Deviation Control compares (near) real-time manufacturing data to predefined thresholds, thus detecting deviations.

The core function Decision Support provides databased support for humans to make informed decisions. By contributing to improving the system, the potential added value of this function is higher than the rather simple deviation control. At the same time, the complexity of the data analysis required to provide decision support is significantly higher than for Deviation Control. The most advanced core function is Autonomous Optimization. Without human interaction, this core function derives decisions based on data analysis and transfers this insight into concrete actions. While the potential value add of Autonomous Optimization is exceptionally high, the requirements are very challenging. As the lack of use cases of fully autonomous prescriptive maintenance demonstrates, data-based Autonomous Optimization is still in its infancy. Generally speaking, a tendency is seen that the potential value added by a core function goes hand in hand with its level of complexity.

Chapter 4.1.4 presents a description of each DBA along with a use case example from literature and their requirements. Based on the requirements discussed for the individual DBAs, requirements are collected and grouped into three requirements categories, comprising 10 key requirements (see Table 17). The three key requirement categories are Technical infrastructure (1), Data availability (2), and Know-how (3).

Technical infrastructure comprises the physical and IT infrastructure needed for data collection, transfer, processing, and protection. Data availability includes access to three distinct kinds of data, namely, historical data, (near) real-time data from the manufacturing system, and central data from the MES or ERP system. The requirement category Know-how addresses the needed expertise of the IT system, manufacturing domain expertise, and DBA specific expertise.

By comparing the requirements of all DBAs, three conclusions can be made. First, reviewed literature has a clear focus on technological aspects of DBAs, hence the majority of requirements discussed are technological in nature with only a few addressing employee capabilities and none addressing organizational capabilities. Second, the DBAs reveal different levels of complexity, which is reflected by different requirements. While some DBAs, such as Track and Trace, only need an appropriate technical infrastructure of tags and readers, more sophisticated DBAs, such as Predictive Maintenance, strongly depend on manufacturing domain knowledge and expert knowledge specific to the DBA. Third, the more complex DBAs can build on the data foundation created by the rather basic DBAs (e.g., Condition Monitoring can provide the data needed for Predictive Maintenance).

In summary, chapter 4.1 provides a sound and literature backed basis to answer the first SRQ: Which data-based application exist in manufacturing and what are their objectives? The DBA use cases demonstrate a high potential to increase the effectiveness and efficiency of a manufacturing system.

Chapter 5 will contrast the potentials identified in this chapter to the actual status quo of data utilization in three industry companies. Furthermore, by consolidating the requirements originating from the discussion of the individual DBAs in chapter 4.1.4, chapter 4.1 also addresses the second SRQ: What are key enablers to apply data-based applications?

However, due to the technical focus of the papers reviewed in this chapter, the identified requirements are also mainly technical. Chapter 5 will complement the collection of critical enablers by focusing on organizational requirements and required employee capabilities in qualitative case studies and expert interviews.

4.2 Impact of Data-based Applications on Lean Practices

This chapter builds on the findings of chapter 4.1. It addresses the third SRQ by evaluating how DBAs impact the implementation of lean practices. This chapter is structured as follows: chapter 4.2.1 describes the methodological approach. Chapter 4.2.2 presents the DBA – Lean Practice Impact Matrix, which summarizes and visualizes potential impacts of DBAs on lean practices. Chapter 0 provides more details on each DBA – lean practice impact by describing the kind and degree of impact. Finally, chapter 4.2.4 provides a summary of the most impactful DBA – lean practice relations.

4.2.1 Methodology

The basic motivation for researching the impact of DBAs on lean practices originated from the result of the Lean2020 study (see chapter 3). Participants reported a large potential of Big Data utilization for lean. Furthermore, mature and successful lean companies show a higher level of data collection and data utilization. This chapter strives to address the question of how data can support lean in a systematic way. According to Mayr et al. (2018, p. 623), current publications address the impact of SM technology on lean on a rather general level, while missing the link to a particular lean practice. Taking this remark into consideration, this research chooses lean practices as appropriate points of reference to evaluate the impact of data utilization on LM.

The selected methodology of a pairwise evaluation of the impact of the DBAs on lean practices is inspired by a dissertation from Gerberich (2011). He investigated the interaction of MES functionalities and lean elements in the automotive industry and developed a “Lean-MES-Interdependency Matrix” (Gerberich, 2011, p. 226). The matrix visualized whether an MES function and a lean element are supporting each other to achieve an objective (positive impact), contradicting, or substituting each other (negative impact) or reveal no impact.

Following this methodology, chapter 4.2.2 presents a DBA – Lean Practice Impact Matrix. Therefore, every combination of the 10 lean practices introduced in chapter 2.2.3 and the 14 DBAs identified in chapter 4.1 is assessed individually as to whether the DBA has a positive impact, a negative impact, or no impact on the respective lean practice.

The assessment followed a four-stage procedure, visualized in Figure 16. The first stage Literature Research comprises the identification of literature that addresses the conjunction of lean in general, not limited to lean practices, and the use of advanced technologies. The consulted literature is depicted in Table 18. Impacts of data utilization on lean practices documented in the literature were directly used as input for the DBA – Lean Practice Impact Matrix (see Table 19). Impact evaluations that are directly based on literature input include references to the respective literature in the DBA – lean practice impact discussion (see 4.2).

The reviewed literature did only provide statements on some DBA – lean practice impacts (e.g., the DBA *Predictive Maintenance* supports the same objective as the lean practice Preventive Maintenance), while many combinations would remain blank after stage 1.

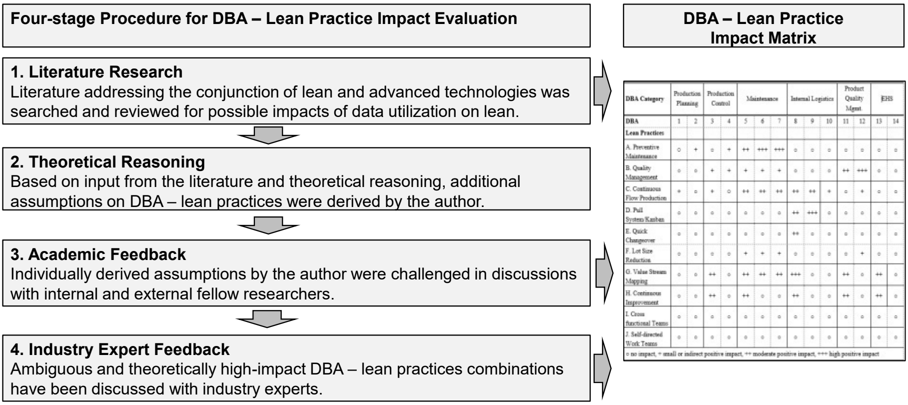

Figure 16: Four-stage procedure: DBA – lean practice impact evaluation

Hence, the second stage *Theoretical Reasoning* goes beyond the evidence provided by literature. By theoretical reasoning additional DBA – lean practice impacts assumptions have been derived by the author. However, as these findings were not grounded in literature, conducting a sense check was reasonable prior to including the assumptions into the DBA – Lean Practice Impact Matrix. The sense check was conducted in the third and fourth stages.

In the third stage, the personal assumptions were challenged in internal discussions with research associates at the ITEM-HSG as well as in discussions with a fellow researcher at the chair of Production and Operations Management at the ETH Zurich. The adapted findings were included in the matrix with two exceptions. If a specific DBA – lean practice impact was either ambiguous (e.g., does DBA *Track and Trace* have a positive or negative impact on lean practice Value Stream Mapping*?*) or potentially highly impactful (e.g., the DBA *Predictive Maintenance* on the Lean Practice Continuous Flow), the DBA – lean practice combination was discussed with practioners from the industry as part of case study interviews[21] in stage four. The restriction on controversial and high impact combinations was necessary due to the limited resources of the case study partners.

Table 18 provides an overview of the consulted literature in the first stage. If only a particular chapter of a book is of relevance, the reference includes the respective page.

Table 18: Input literature to assess the impact of DBAs on lean practices

Focus	Reference
Lean and IT systems	Bell (2006), Clegg and Powell (2013), Gerberich (2011), Kletti (2015) Maguire (2016), Powell et al. (2013), Ward and Zhou (2006)
Lean and SM / Industry 4.0	Bertagnolli (2018, p. 192), Bick (2014), Buer et al. (2018), Dombrowski et al. (2017), Karre, Hammer, Kleindienst, and Ramsauer (2017), Kieviet (2016, p. 42) , Mayr et al. (2018), Meissner et al. (2018), Metternich et al. (2017), Mrugalska and Wyrwicka (2017), Rüttimann and Stöckli (2016), (Sanders et al., 2016), Wagner et al. (2017)
Lean automation	Hedelind and Jackson (2011), Kolberg et al. (2016), Kolberg and Zühlke (2015)
Other	Hicks (2007), Uriarte, Ng, and Moris (2018), Rafique et al. (2016)

As shown in Table 18, two main clusters can be identified. Several authors discuss the role of IT systems in LM, including the use of central IT systems as ERP (Clegg & Powell, 2013), MES (Gerberich, 2011), and IT as a driver for CI (Bell, 2006). The second cluster is by far the most comprehensive one and comprises authors addressing the interaction of lean and SM from several perspectives. These publications cover the two dimensions of lean as an enabler of SM and SM as a toolbox to advance lean (see chapter 2.4). A small cluster is found around lean automation. The cluster *Others* include contributions on the application of lean thinking to information management (Hicks, 2007), supporting lean with simulations (Uriarte et al., 2018), and the application of RFID to mitigate barriers of LM (Rafique et al., 2016).

[21] For details on the interview partner and the case study companies, see chapter 5.2.1

4.2.2 DBA – Lean Practice Impact Matrix

4.2.2.1 Evaluation Criteria

The matrix differentiates three levels of positive impact. The impact of a DBA on a lean principle is considered as high (+++) if the DBA enables the lean practice to overcome an existing barrier or the DBA and the lean practice are fully complementary regarding achieving a common objective. Furthermore, the impact needs to be both theory-based and confirmed as significant by practitioners.

The impact of a DBA on a lean principle is considered as low (+) if the lean principle is supported only to a small degree or the impact is only indirect. Indirect impact refers to a situation, where a DBA has an objective that is distinct from the objective of the lean practice and the support is rather a byproduct. For example, the DBA *Predictive Maintenance* increases equipment health. Equipment health, in turn, is positively linked to product quality. Therefore, maintenance has an indirect positive impact on the lean principle of Quality Management. Impacts between high and low are considered as moderate (++). No impact is marked with (○). The possibility of negative impacts have been considered, and scenarios of negative impacts of a DBA on a lean practice are also discussed in the next chapter.

At this point, it has to be noted that the evaluation of the impact levels is subject to a certain degree of subjectivity of the authors and involved partners from academia and industry. In a different environment, other evaluation judgments are conceivable. However, discussions with researchers and practitioners have revealed a fairly good match regarding the DBA – lean practice impact evaluation. Although there have been discussions about the extent of the impact, there has usually been an agreement about the basic impact of a DBA on a lean practice.

4.2.2.2 Overview

Table 19 shows the DBA – Lean Practice Impact Matrix, including six framed clusters of a high positive impact of DBAs on lean practices. All identified DBA – lean practice combinations that have been found to have an impact are discussed in the following chapter, including a rationale for the respective evaluation of the degree of impact. In addition, the six clusters of high positive impact are summarized in chapter 4.2.4.

The DBA – Lean Practice Impact Matrix provides an overview of all DBA – lean practice combinations discussed in this dissertation. Reading example: The DBA *Production Scheduling* (2) has a small positive impact on the lean practice Preventive Maintenance (A).

Table 19: DBA – Lean Practice Impact Matrix

DBA Category	Production Planning		Production Control		Maintenance			Internal Logistics			Product Quality Mgmt.		EHS	
DBA / **Lean Practices**	1	2	3	4	5	6	7	8	9	10	11	12	13	14
A. Preventive Maintenance	○	+	○	+	++	+++	+++	○	○	○	○	○	○	○
B. Quality Management	○	○	+	+	+	+	+	○	+	○	++	+++	○	○
C. Continuous Flow Production	+	○	+	○	++	++	++	++	++	+	○	+	○	○
D. Pull/Kanban	+/-	○	○	○	○	○	○	++	+++	○	○	○	○	○
E. Quick Changeover	○	○	○	○	○	○	○	++	○	○	○	○	○	○
F. Lot Size Reduction	○	○	○	○	+	+	+	○	○	○	○	+	○	○
G. Value Stream Mapping	○	○	++	++	++	++	++	+++	○	○	++	○	++	○
H. Continuous Improvement	○	○	++	+++	++	++	++	++	○	○	++	○	++	○
I. Cross-functional Teams	○	○	○	○	○	○	○	○	○	○	○	○	○	○
J. Self-directed Work Teams	○	○	○	○	○	○	○	○	○	○	○	○	○	○

- negative impact, ○ no impact,
+ small or indirect positive impact, ++ moderate positive impact, +++ high positive impact

1. Layout Planning, 2. Production Scheduling, 3. Real-time Control, 4. System Performance Measurement, 5. Condition Monitoring, 6. Predictive Maintenance, 7. Prescriptive Maintenance, 8. Track and Trace, 9. Material Flow Management, 10. Inventory Management, 11. Product Quality Monitoring, 12. Product Quality Improvement, 13. Energy Monitoring, 14. Environment Monitoring

4.2.3 DBA – Lean Practice Impact Discussion

This chapter is structured along the four lean practice categories TPM, TQM, JIT, and EMS as introduced in chapter 2.2.3. It elaborates the DBA – lean practice relations visualized in Table 19 in more detail. If supporting evidence is available, the argumentation of the impact evaluation is supported by references to the literature shown in Table 18.

No reference indicates that the impact evaluation is based on logical reasoning and personal assumption. The assumptions are presented in its final form, which means after the initial assumption was refined in discussions with other researches and practitioners.

4.2.3.1 TPM Lean Practice

4.2.3.1.1 Preventive Maintenance

The lean practice Preventive Maintenance aims to minimize unplanned production interruptions due to unexpected machine breakdowns by performing maintenance tasks preventively. Thus, the practice is strongly supported by the DBAs of the category Maintenance (DBA 5-7 in Table 19).

4.2.3.1.1.1 Applications with High and Moderate Impact

DBA *Condition Monitoring* (Number 5 in Table 19)

The application *Condition Monitoring* assesses the health and degradation status based on machine data to detect anomalies early and thus enable maintenance employees to initiate maintenance activities before the machine breaks down (Sanders et al., 2016, p. 825). Thus it increases equipment availability and avoids consequential damage of products due to broken machines (Mayr et al., 2018, p. 625).

Condition Monitoring provides a data-based foundation for better maintenance plans and therefore supports the lean practice preventive maintenance. Impact evaluation: moderate.

DBA *Predictive Maintenance* (6)

The application *Predictive Maintenance* not only assesses the equipment condition but also forecasts expected future conditions. Predictive analytics increases the accuracy of remaining useful lifetime prognosis and thus, enables the definition of maintenance plans that provide an optimal balance between ensuring equipment uptime and reducing unnecessary maintenance cost. (Kieviet, 2016, p. 55; Mayr et al., 2018, p. 625; Sanders et al., 2016, p. 825).

Predictive Maintenance provides powerful support for the lean practice Preventive Maintenance by providing predicted machine conditions as a foundation for well-informed maintenance decisions. The DBA is fully complementary regarding the objectives of the lean practice Preventive Maintenance. Impact evaluation: high.

DBA *Prescriptive Maintenance* (7)

The argumentation for *Prescriptive Maintenance* is equivalent to the *Predictive Maintenance*.

4.2.3.1.1.2 Applications with Low Impact

DBA *System Performance Measurement (4)*

The DBA *System Performance Measurement* provides automated performance metrics. Monitoring the OEE and using the metric for benchmarking purposes allows employees to identify OEE improvement potentials of monitored equipment (Gerberich, 2011). Thereby, the application also provides indirect support for Preventive Maintenance: Impact evaluation: low.

DBA *Production Scheduling (2)*

The expected intensity of use of a specific machine for the next time period can be derived from the *Production Scheduling* plan. This information can be used to deploy limited maintenance resources more effectively according to the importance of the machine to execute the next production plan. Impact evaluation: low.

4.2.3.1.1.3 Conclusion

The lean practice Preventive Maintenance benefits from different application of data utilization and existing limitation of current Preventive Maintenance practices can be mitigated. By using machine condition data and prediction of future conditions, companies can better determine the optimal time for maintenance. Thereby, a better balance of equipment availability and maintenance costs is achieved.

4.2.3.2 TQM Lean Practice

4.2.3.2.1 QUALITY MANAGEMENT

The lean practice Quality Management is a part of the broader lean practice Total Quality Management (see chapter 2.2.3). In this context, Quality Management focuses on product quality.

4.2.3.2.1.1 Applications with High and Moderate Positive Impact

DBA *Quality Monitoring* (11)

The DBA aims to identify and reject incorrect parts before they are passed on to the next process step or delivered to the customer. Sensors and cameras test whether the characteristics of the product (e.g., thickness and surface roughness) are in accordance with the specification (Tao et al., 2018, p. 9). If a product does not meet the requirement, it is sorted out or used to identify the cause of the deviation (Wuest et al., 2014, p. 1169).

Industry feedback: an interviewee (Company B) confirms the high potential of automated quality monitoring. Accordingly, quality monitoring systems are for many parts more accurate than a human quality controller, faster, and in the long run less expensive. However, he also points out that even modern and highly expensive camera systems are not able to perform a 100 percent accurate visual quality control. Nevertheless, automated quality monitoring is a trend and implemented more and more.

By detecting defect products, the DBA directly contributes to the objective of Quality Management. However, as the application *Quality Monitoring* is only reactive and does not include quality improvement, the impact is only considered as moderate.

DBA *Quality Improvement* (12)

Gerberich (2011, p. 244) has evaluated the relationship between MES functionalities and lean elements. Accordingly, data collection, as part of the MES system, strongly supports the identification of root causes of product quality problems. By providing relevant data, lean elements concerned with quality improvements, such as the PDCA cycle or Ishikawa analysis, can be applied more effectively and efficiently (Gerberich, 2011, p. 244). By monitoring product quality, influencing factors for product quality can be identified and optimized (Künzel, 2016, p. 59). *Quality Improvement* goes beyond correcting or sorting out defect parts, but uses historical failure data to learn from in order to find and mitigate systematic root causes of quality issues.

The overall objective of Quality Management, to provide high-quality and defect-free products to customers, is directly supported by the DBA *Quality Improvement.* Impact evaluation: high.

4.2.3.2.1.2 Applications with Low Positive Impact

The positive impacts of the three maintenance DBAs—*Condition Monitoring (6)*, *Predictive Maintenance (7)*, and *Prescriptive Maintenance* (8)—on product quality are apparent. However, production equipment in good condition produces high-quality products. Tools worn down will sooner or later result in lower quality of the products. As all three DBAs contribute to higher equipment quality, they indirectly influence product quality positively. Impact evaluation: low.

The application *Real-time Control (3)* informs shop floor personnel in case of anomalies in the manufacturing process. As these anomalies may affect product quality, reducing the time of non-conforming process behavior decreases the likelihood of products of bad quality. For instance, Sanders et al. (2016, p. 825) highlight the positive effect of statistical process control on product quality.

Real-time Control supports the practice Quality Management indirectly. Impact evaluation: low.

System Performance Measurement provides quality performance metrics such as scrap rate, which are useful to identify negative trends in terms of quality. Detecting declining quality may serve as a starting point for quality improvement projects. Impact evaluation: low.

The DBA *Track and Trace* includes the use of RFID tags to identify products or containers. These tags cannot only be used for tracing the product but also to provide product-specific information, including details about the operations to be done on them (Bell, 2006, p. 314). This may include supporting data for manual operations (Sanders et al., 2016, p. 825), such as ensuring the correct order of construction. As the support is only indirect, the impact of *Track and Trace* on the lean principle Quality Management is rated as low.

4.2.3.2.1.3 Conclusion

The lean practice Quality Management is significantly supported by the DBAs that are designed to detect defect products and to increase product quality. Well-maintained production equipment supports product quality indirectly by reducing negative impacts due to worn down tools.

4.2.3.3 JIT Lean Practice

4.2.3.3.1 Continuous Flow Production

In Continuous Flow Production, the material or product flows through the value-adding process steps without interruptions and waiting times between the process steps.

4.2.3.3.1.1 Applications with High and Moderate Positive Impact

DBAs *Condition Monitoring* (5), *Predictive Maintenance* (6) and *Prescriptive Maintenance* (7)

Continuous Flow Production requires a high level of equipment availability (Womack & Jones, 2003, pp. 60–61). Therefore, the DBAs ensuring high equipment availability are likely to have a positive impact on Continuous Flow Production.

As discussed above, the three DBAs assigned to the group of maintenance DBAs increases machine availability. Less machine downtime, especially less unexpected machine downtime, increases process stability. Stability is favorable for continuous flow (Mayr et al., 2018, p. 624; Wagner et al., 2017, p. 128). The impact of the maintenance DBAs on Continuous Flow Production is rated as moderate.

Industry feedback: the interview partner of Company B shares the theoretical argumentation and points out that *Predictive Maintenance* may increase the OEE and stability of a machine. However, stability is only one requirement of flow among others (e.g., timely material supply), and the effect of good maintenance on flow is therefore only evaluated to be moderate.

DBA *Track and Trace* (8)

Sanders et al. (2016, p. 823) have identified errors in inventory counting and capacity shortages as a significant cause for disruption in material flow. The application *Track and Trace* employs RFID technology and ensures an error-free inventory status realized by real-time and exact tracking of inventory. Therefore, the application Track and Trace supports Continuous Flow Production. Impact evaluation: moderate.

DBA Material Flow Management (9)

Based on real-time demand and material availability, the DBA *Material Flow Management* may use a mathematical optimization model to develop an optimized material distribution plan. The plan minimizes interruptions and waiting times, enabling a continuous material flow (Sanders et al., 2016, p. 823). The data-based optimization model can also optimize path planning for AGVs, thus supporting JIT delivery (Mayr et al., 2018, p. 624). Dombrowski et al. (2017, p. 1065) have analyzed 260 use cases of industry 4.0 to get an overview of the correlations between LPSs and industry 4.0. Accordingly, the combination of RFID tags and big data analysis for intelligent material flow planning supports the flow principle. The impact of the DBA *Material Flow Management on* the lean practice Continuous Flow Production is rated as moderate.

4.2.3.3.1.2 Applications with Low Positive Impact

DBA *Layout Planning* (1)

Continuous Flow Production requires a layout, which facilitates the flow of products from one process step to the next one without high transportation effort. The conformance of the layout plan with the continuous flow requirements can be defined as a criterion for the optimization algorithm of the layout planning application. Thus, a smartly defined layout planning tool can facilitate Continuous Flow Production. Nevertheless, the impact is evaluated as low, as designing a continuous flow ready layout plan does not necessarily require data-based *Layout Planning*.

DBA *Real-time Control* (3)

The application reduces equipment downtime, but only after a problem has already occurred. It minimizes the time in which the uninterrupted flow of material is inhibited due to production stops. Impact evaluation: low.

DBA *Inventory Management* (10)

While the application *Track and Trace* focuses on having a high transparency on the material flow, *Inventory Management* tracks the current amount of inventory in the warehouse. The status of the warehouse inventory, however, is less important for Continuous Flow Production as the status of WIP inventory. Thus, the impact of *Inventory Management* on Continuous Flow Production is only rated as low.

DBA *Product Quality Improvement* (12)

The DBA uses historical data from defective products to systematically identify and permanently eliminate the root causes of defects. Continuous Flow Production requires not only stable equipment but also flawless products (Womack & Jones, 2003, p. 61). Thus, the lean principle is supported by the DBA *Quality Improvement* indirectly. Impact evaluation: low.

4.2.3.3.1.3 Conclusion

Continuous Flow Production is enabled by several factors; among these are stable equipment and reliable material availability. The stability of production equipment, on one hand, is improved by the applications of the category *Maintenance.* Material availability, on the other hand, is ensured by a smart material flow organized by the application *Material Flow Management.* In conclusion, these DBAs can significantly contribute to the reduction of barriers to Continuous Flow Production.

4.2.3.3.2 PULL/KANBAN

4.2.3.3.2.1 Applications with High and Moderate Positive Impact

DBA *Track and Trace* (8)

Track and Trace uses RFID and other technologies to monitor the amount and location of production material. Thus, the current material availability can be monitored in real-time (Sanders et al., 2016, p. 822). If the material stock falls below a specified minimum threshold value, material replenishment is triggered automatically. *Track and Trace* facilitates the monitoring of material in supermarkets and therefore supports the lean principle Pull/Kanban. Impact evaluation: moderate.

DBA *Material Flow Management* (9)

The application *Material Flow Management* builds on track and trace and other data to organize an intelligent material flow. For instance, e-Kanban systems recognize empty bins automatically and trigger replenishment. Instead of cards, the signal is transmitted wirelessly. The concept of e-Kanban is very popular and often presented as an example of enhancing lean by means of industry 4.0 technologies (Bertagnolli, 2018, p. 193; Clegg & Powell, 2013, p. 1499; Kolberg et al., 2016, p. 2851; Sanders et al., 2016, p. 822). E-Kanban mitigates some limitation of traditional Kanban (Clegg & Powell, 2013, p. 1499) and therefore provides a direct and strong support of the lean practice Pull/Kanban. If AGVs are involved in material handling, optimization algorithms can be used for intelligent route planning, based on current demand (Bertagnolli, 2018, p. 194). In summary, the DBA provides a strong positive impact on the lean practice *Pull/Kanban*.

4.2.3.3.2.2 Applications with Ambiguous Impact

DBA *Production Scheduling* (2)

The combination of IT-supported planning and lean has received considerable attention from scholars. It serves as the perfect example of the contradiction between lean and IT technology (Maguire, 2016, p. 32). The precursors of today's ERP systems operated according to the *push* principle, while lean promotes the *pull* principle (Maguire, 2016, p. 34). Taking this contradiction for granted implies a perfect substitution of the lean principle Pull/Kanban by the DBA *Production Scheduling*. Consequently, the impact of the DBA on the lean practice would be strongly negative. However, as Clegg and Powell (2013, p. 1502) point out, modern ERP systems are by now also able to support the pull approach and combine intelligent planning and the execution of the pull lean practices. Gerberich (2011, p. 234) notes the possibility of a hybrid system that uses both approaches, thus creating a so-called *push-pull interface.*

Industry feedback: interviews with two lean managers have revealed different approaches to integrate central production planning and pull in the industry. A lean manager of a major automotive supplier (Company B) explained that in their factory a central planning system plans the material flow between the individual production process steps. By eliminating the supermarkets between the process steps, WIP inventory is reduced. The concept of pull, however, is still in place on a higher level as the whole production process is triggered by a customer order.

An interview partner fo Company A, working as an specialist for connected logistics, does not see a conflict between IT-supported production planning and pull, as his company uses both approaches simultaneously. The system is planning bottleneck processes centrally to ensure material availability. The rest of the material flow is organized according to the pull principle, using an electronic Kanban system. He argues that the smart combination of IT-supported material flow planning and the e-Kanban system yields the best results in terms of material availability and low WIP inventory.

Taking all aspects into consideration, neither a strictly positive nor a negative impact of the DBA *Production Scheduling* on the lean practice Pull/Kanban is apparent.

4.2.3.3.2.3 Conclusion

In theory and literature, the pull principle is threated to be substituted by the push principle due to central production and material flow planning systems. However, developments of those planning systems now also support pull. Discussions with lean responsibles revealed that the industry currently uses both approaches simultaneously, resulting in a partial substitution of the lean practice Pull/Kanban by central planning.

A small positive impact on the lean practice Pull/Kanban can be expected from the increasing use of Track and Trace to monitor material availability in real-time. A strong positive impact on the lean practice Pull/Kanban originates from the implementation of e-Kanban, which can be assigned to the DBA *Material Flow Management*. E-Kanban is faster, flexible, and cards cannot get lost (Clegg & Powell, 2013, p. 1499) and thus mitigates some of the weaknesses of traditional Kanban.

4.2.3.3.3 QUICK CHANGEOVER

The lean practice Quick Changeover reduces the time for changeovers to allow small lots being produced economically. Reviewed literature provides only one example of support for this lean practice.

4.2.3.3.3.1 Applications with High and Moderate Positive Impact

DBA *Track and Trace* (8)

Comparably frequently discussed is the ability of RFID tags to carry product-specific data. As the product approaches the respective machine, the operations to be performed are transmitted to an RFID receiver. The machine can change to the required tools and settings to fit the requirement of the product before the product arrives, thus reducing setup time substantially (Sanders et al., 2016, pp. 823–824). Impact evaluation: moderate.

4.2.3.3.3.2 Conclusion

Quick Changeovers are getting increasingly important for producing different product variants without spending much time on changeovers. Product-specific data on RFID tags allow machines to adjust to the requirements of the operation to be performed before the product has arrived at the machine. No further DBAs supporting this lean practice have been found.

4.2.3.3.4 LOT SIZE REDUCTION

For the lean practice Lot Size Reduction, no reference of a possible impact of any of the 14 DBAs has been found in the reviewed literature. However, based on the assumption that small lot sizes imply little WIP and security stock, two possible impact scenarios can be imagined.

4.2.3.3.4.1 Applications with Low Positive Impact

Maintenance DBAs (6-8)

If only little WIP inventory and security stock are held, it is essential to have reliable equipment to ensure the timely supply of needed parts. The impact of the maintenance DBAs on the lean practice Lot Size Reduction is minor and indirect. Impact evaluation: low.

DBA *Product Quality Improvement* (12)

Building on the same argument, little WIP and security stocks require defect-free parts for further processing to avoid production stops. The impact is minor and indirect. Impact evaluation: low.

4.2.3.3.4.2 Conclusion

By assuring machine stability to produce needed parts quickly and assuring low defect rates, DBAs can contribute to the ability to work with small lot sizes. The impact, however, is limited and in conclusion rated low.

4.2.3.3.5 VALUE STREAM MAPPING

The lean practice Value Stream Mapping is used to visualize the current production process. Increasing accuracy and reducing the effort of value stream analysis by integrating production data into the VSM process is a frequently discussed approach to improve lean through data utilization (Buer et al., 2018, pp. 2930–2931; Mrugalska & Wyrwicka, 2017, p. 471; Prinz et al., 2018, p. 23; Tantik & Anderl, 2016, p. 208).

J. C. Chen and Chen (2014, p. 839) propose a system to create a value stream map automatically, using performance monitoring data. Creating the VSM automatically reduces errors and human effort. Furthermore, by integrating real-world manufacturing data, the value stream map represents the reality more accurately and supports supervisors to make more informed decisions on the shop floor. Particularly for companies manufacturing a high variety of products, IT-supported value stream mapping is a promising tool to adequately reflect the minor differences in the process between the different types of product. This approach is also referred to as VSM4.0 (Buer et al., 2018, p. 2930). Tantik and Anderl (2016, p. 208) even propose to assign the current value to each product, based on the current position in the manufacturing process. Following Mayr et al. (2018, p. 625), VSM4.0 increases the transparency in the value chain and facilitates the identification of waste, thus enabling a lean value creation.

Industry Feedback

Due to the potentially high impact of VSM4.0 on the lean practice Value Stream Mapping, the presented suggestion of VSM4.0 was discussed with representatives of the three case companies.

In general, the interviewees appreciate the idea of integrating real-time manufacturing data into the value stream mapping process. An interviewee of Company A argues that especially in regard to the increasing number of variants with possibly distinct value streams, standard value stream mapping is very time-consuming. Manufacturing data are considered to be valuable to uncover dynamic changes in the process. For instance, bottlenecks are not static and may change over time and from product to product. All three interviewees, however, agree that the basic idea of Value Stream Mapping will get lost if *go (to the process) and see* is replaced by the remote extraction of data from the IT system. The objectives of VSM includes gaining a personal understanding of the actual process, which is impossible without being on the shop floor. The interviewees point out that without understanding the underlying process, the risk of data misinterpretation rises considerably. Consequently, they opt for a combination of traditional VSM and the usage of real-time manufacturing data. By bringing the best of the two worlds together, process understanding is ensured by *go and see* while manufacturing data is integrated to reduce human effort and increase the accuracy of the value stream map. One partner puts it like this: “Go and see plus measure is always better than measure and interpreting data at the computer.”

4.2.3.3.5.1 Applications with High and Moderate Positive Impact

Data required for VSM4.0 is collected by several DBAs. *Real-time Control* tracks anomalies in the production process such as machine stops, *Condition Monitoring* provides data on the current equipment health.[22] Performance indicators that, for instance, measure the average lead time for a large sample is provided by the application *System Performance Measurement*. *Product Quality Monitoring* provides data on defect and rejected parts. *Energy Monitoring* reveals potentials for energy waste reduction. By providing the required data, these DBAs support the lean practice VSM. Impact evaluation: moderate.

Most relevant for measuring the actual value stream, however, are track and trace data. This data helps to identify bottlenecks as well as situations of long waiting times and high inventory. Hence, the impact of Track and Trace is rated as high.

[22] The DBAs Preventive Maintenance and Predictive Maintenance also have the functionality of condition monitoring, hence have the same potential to support the lean practice VSM. The additional functionalitities of the more advanded maintenance DBAs, however, have no additional value. Therefore, the rating is equal.

4.2.3.3.5.2 Conclusion

In conclusion, the lean practice VSM is likely to benefit substantially from integrating manufacturing data into the value stream map. VSM can be conducted more accurately and reflect different value streams of different product variants. Updates of the value stream map require less human effort and can also reflect changes over time. However, manufacturing data extracted from IT systems is only valuable as a complementary source of information and must not replace the presence on the shop floor.

4.2.3.4 EMS Lean Practice

The category Effective Management System is concerned with the objective to motivate and align people to work for a common goal. From the 10 lean practices presented in chapter 2.2.3, the following lean practices are assigned to this group: *Continuous Improvement*, *Cross-Functional Teams*, and *Self-directed Work Teams.* Like *Value Stream Mapping*, *Continuous Improvement* will benefit from the availability of more and more accurate data. For the other human-centered lean practices—*Cross-Functional Teams* and *Self-directed Work Teams*—examples for data-based support have neither been found in the literature nor have been derived theoretically.

4.2.3.4.1 CONTINUOUS IMPROVEMENT

Manufacturing data increases the transparency of the production. This transparency is used to identify opportunities for improvement, which is the starting point of CI. According to Bell (2006, p. 36), delivering the right information, in the right format, to the right place, and at the right time is “a powerful tool for continuous improvement.” The CI tools PDCA and DMAIC both comprise a planning (PDCA) or an analyzing (DMAIC) phase that rely on accurate data (Sokovic et al., 2010, p. 480). Hence, particular CI tools can be applied more effectively with access to the right kind of data at the right time. Several researchers cited in this work support the assumption that CI can significantly benefit from manufacturing data (Gerberich, 2011, p. 79; Meissner et al., 2018, p. 83; Prinz et al., 2018, p. 23; Wagner et al., 2017, p. 128).

4.2.3.4.1.1 Applications with High and Moderate Positive Impact

Similar to *Value Stream Mapping,* it is difficult to assess the impact of the 14 DBAs on CI individually, as the required data depends on the CI approach and its objective. Therefore, the DBAs most relevant for monitoring and collecting data, including *Real-time Control, Condition Monitoring, Track and Trace, Product Quality Monitoring*, and *Energy Monitoring* are equally considered as facilitators of CI. By providing relevant data in the right quality, CI tools can be applied more effectively (Gerberich, 2011, p. 244). The listed DBAs, therefore, provide a direct support for the lean practice CI. Impact evaluation: moderate.

DBA *System Performance Measurement* (4)

The DBA System Performance Measurement is especially valuable to support CI. While the DBAs listed above can serve as data provider to better understand a given problem, the DBA System Performance Measurement has an additional, highly important function. By calculating and visualizing KPIs, the DBA allows for the ability to track trends and do benchmarking against reference values. Hence, the DBA supports employees in detecting weaknesses within the production system and thus serve as a starting point for diverse CI activities addressing these weaknesses. Impact evaluation: high.

4.2.3.4.1.2 Conclusion

The lean practice Continuous Improvement comprises several aspects; among these are encouraging employees to think about improvement activities constantly as well as providing the right tools and information needed for problem-solving. DBAs such as the DBA *System Performance Measurement* increases transparency about the production system and facilitates the identification of hidden problems. Depending on the actual problem, different data may be required to identify the problem's root cause. The need for accurate up-to-date data can be met by those DBAs, which include a data monitoring function. Hence, DBAs can support CI in two ways: detecting problems and providing the data for an effective root-cause analysis.

4.2.3.4.2 CROSS-FUNCTIONAL TEAMS AND SELF-DIRECTED WORK TEAMS

The literature review did not reveal an example of data-based support of any DBA for the two human-centered lean practices Cross-Functional Teams and Self-directed Work Teams. Also, the author of this dissertation did not find a scenario in which the two lean practices unambiguously benefit from the availability of manufacturing data. Cross-Functional Teams and Self-directed Work Teams are concerned with organizing the work but, unlike the other DBAs, are not inherently linked to the product or the production process. Without this link, it seems reasonable to assume that product and process-related manufacturing data is generally of little relevance for these two lean practices.

The direction of support is even reversed in this case. As seen in the DBA requirement analysis in chapter 4.1.6, several DBAs require different skills, comprising IT-know-how, manufacturing domain expertise, and DBA specific expertise. This broad collection of required skills is seldom found within one function and hence calls for cross-functional teams comprising several skill sets.

4.2.4 Summary Impact of DBAs on Lean Practices

The pairwise analysis has revealed the potential for broad support for lean practices by utilizing manufacturing data. Table 19 indicates that at least six out of 10 considered lean practices are likely to be highly supported by one or more DBAs.

First, the detailed evaluation in chapter 0 shows a strong positive impact of *Predictive Maintenance* on the lean practice Preventive Maintenance. *Predictive Maintenance* is a well-established application of data utilization in manufacturing. It is not only discussed frequently in academic articles as an example of data utilization in manufacturing but is also seen by practitioners as a promising approach to increase equipment availability and process stability. All interviewed companies report having *Predictive Maintenance,* at least to some extent, already in use. A barrier for a broader implementation at the moment are high costs for *Predictive Maintenance,* which in many cases still exceed the costs for changing spare parts routinely by following a fixed plan. However, there is little doubt that *Predictive Maintenance* will be increasingly relevant and will eventually advance to the new standard of maintenance.

Second, the application *Quality Improvement* has a strong positive impact on the lean principle Quality Management. Literature suggests that by monitoring and collecting root cause analysis can be conducted faster and more systematically and thus more effectively. To this end, ML is a promising technique to classify and detect failures but also to identify their defect root causes.

In general, data-based product quality improvement is perceived as highly relevant by the case companies. The partner companies follow different approaches to use data for quality improvements. Company A has a dedicated quality department which has access to product data, machine data, and track and trace data. In this department, quality experts conduct data supported root-cause analysis and provide feedback to manufacturing and R&D. In Company B, failure data is used for root-cause analysis primarily as part of shop floor management meetings.

Third, Continuous Flow Production is supported from two perspectives. On one hand, maintenance DBAs decrease the likelihood of unexpected machine breakdowns and therefore impact process stability positively. Process stability, in turn, is essential for Continuous Flow Production; consequently, the lean practice is supported by purposeful maintenance activities. On the other hand, Continuous Flow Production depends on the timely availability of material. Tracking the current status of material availability with Track and Trace and based on the data, derive an intelligent material distribution plan supports the reduction of interruptions and waiting times, thus enabling a continuous material flow.

Fourth, the Pull System is strongly supported by the DBA *Material Flow Management.* Very popular in academic literature is the concept of e-Kanban. Scholars have identified several advantages of e-Kanban compared to traditional Kanban including faster transmission of signals, no lost cards, and the ability to adjust the lot size dynamically. The advantages of e-Kanban are confirmed by all interview partners operating a physical production. As a consequence, both manufacturing companies have replaced traditional paper-based Kanban with e-Kanban by now.

Fifth, the lean practice VSM may benefit considerably from integrating real-time manufacturing data into the value stream map. Literature and practitioners acknowledge the potential of real-time data enhanced VSM, also called VSM4.0, to draw a more precise picture of the actual value stream. Especially regarding the trend for increased product variety, one partner argues that the value streams of similar but slightly different products may not be perfectly identical. Manufacturing data can be used to reflect these minor differences as well as dynamic changes in the value stream, for instance, product-specific bottlenecks. Furthermore, having the necessary data available allows conducting VSM regularly with low human effort.

Similarly, CI is supported by access to manufacturing data. CI applies systematic tools such as the DMAIC cycle and Ishikawa diagrams. The rigor of these tools is enhanced by accurate manufacturing data collected by multiple DBAs including *Real-time Control*, *Condition Monitoring*, *Track and Trace*, *Product Quality Monitoring*, and *Energy Monitoring.*

Before a problem can be addressed by CI methods, however, it needs to be detected. By providing automatically calculated KPIs and visualizing their trends the DBA *System Performance Measurement* facilitates the identification of hidden problems within the production system.

The evaluation in chapter, however, has not only indicated several opportunities for lean support by DBAs but has also highlighted some conflicts.

A basic conflict frequently discussed in academic literature is push vs. pull. While LM undoubtedly advocates the pull principle, IT-supported planning and scheduling systems often rely on the push principle. Discussions with two industry partners have shown that, at least to some degree, internal material supply is planned by a central system, thus replacing the lean practice Pull/Kanban for some parts. From a practitioner's view, however, the contradiction described in the literature is of less importance as reality shows that push and pull can be used simultaneously.

The concept of VSM 4.0 has aroused both interest and concern from practitioners. The positive implications have already been discussed above. On the negative side, industry representatives see the risk of substituting the presence on the shop floor for VSM with remote access to manufacturing data. Consistently highlighted is the importance of being “where the action happens” to gain a sound understanding of the real process. Accordingly, process data without process understanding is of little value and bears the risk of data misinterpretation. The basic principle of VSM is described as “go and see.” Therefore, a virtual value stream created by remotely accessible data would contradict the fundamentals of VSM. Nevertheless, a balanced integration of few metrics is desirable as long as it does not replace shop floor presence.

No cases for DBA support for the human-centered lean practices of Cross-Functional Teams and Self-directed Work Teams were found. Taking the DBAs' requirement for cross-functional skills into account, a reverse implication can be argued. Cross-Functional Teams comprising employees with different skill sets are likely to introduce the DBA more effectively than a group of manufacturing engineers or a group of data scientists without complementary expertise.

Lean practices have demonstrated for many years, in many industries, and in many countries worldwide the ability to support the objectives of LM—high quality, low costs, and fast delivery. The previous chapters have found a positive impact of DBAs on several widely established lean principles. By supporting these lean principles, DBAs contribute to the achievement of lean objectives. This chain of effects might be useful for managers to justify investments needed to introduce applications that exploit the potential of manufacturing data.

5 Qualitative Studies

Chapter 4 provided a literature-based foundation to answer the first SRQ ("Which data-based application exist in manufacturing and what are their objectives?") and the third SRQ ("How can data-based applications support lean practices?"). Furthermore, it consolidated key requirements described in the DBA use cases and thus also addressed the second SRQ ("What are key enablers to apply data-based applications?") already. The insights gained from the following qualitative studies are used to complement the findings of chapter 4. Case study research and expert interviews are thereby selected as primary sources of qualitative data

5.1 Case Study

This chapter introduces the methodology of case study research in chapter 5.1.1, describes the case selection in chapter 5.1.2, and discusses the sources for data collection in chapter 5.1.3.

5.1.1 Case Study Methodology

Case study research was selected for the following reasons. Case study research is qualified to investigate problems originating from the practical world because researching real-world problems often results in a high practical relevance of the findings (Gassmann, 1999, p. 11). Also, case study research is suitable for research in areas with little existing knowledge (Voss et al., 2002, p. 198) and to investigate explanatory research questions beginning with *how* and *why* (R. Yin, 2009, p. 36) or *what* (Creswell, 2014, p. 140). While quantitative research is appropriate to test a hypothesis by analyzing large samples, qualitative research aims to obtain generalizable patterns from a small set of cases, thus following the induction principle (Tomczak, 1992, p. 77).

According to Ketokivi and Choi (2014, p. 233), three modes of conducting case research are distinguished (see Figure 17).

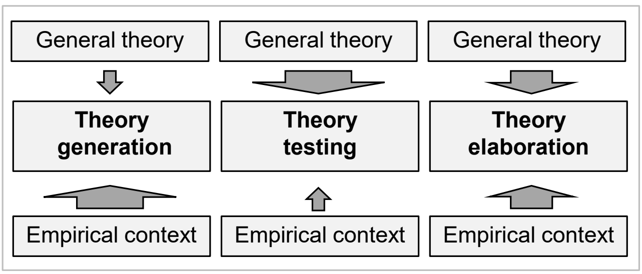

Figure 17: Three modes of conducting case research (Ketokivi & Choi, 2014, p. 233)

All three modes of case research seek to formulate theoretical insight that results from the interaction between an existing general theory offered by literature and the empirical context of the case. Theory generation, testing, and elaboration differ in the relative importance of empirical context and general theory, as represented by different thicknesses of the arrows in Figure 17.

The research at hand is guided by the theory of the TAM, as introduced in chapter 1.5.5. Consequently, the primary objective of qualitative research in this dissertation is not theory generation in the sense of creating a new theory. However, the existing TAM theory is challenged by the findings of the qualitative studies as to whether the theory can explain the major findings from a theoretical perspective. Therefore, the initial objective can be described as theory testing. Compared to theory-generating case studies, theory-testing case studies approach the research with a more a priori theoretical discipline (Ketokivi & Choi, 2014, p. 235). Nevertheless, theory testing may shift to theory elaboration, if evidence for lack of explanatory power is found in the existing general theory (Eisenhardt, 1989, p. 536). Therefore, if such a lack is identified, adapting and extending the TAM theory to increase its explanatory power of the phenomenon at hand is a further objective of this qualitative research.

Eisenhardt (1989) stresses the need for a priori specifications of constructs intended to be researched to increase the accuracy of their measurement. Miles and Huberman (1994) suggest doing this by creating a conceptual research framework, as introduced in chapter 1.5.4. Clearly defined research questions are essential to define a clear scope, which in turn supports the selection and collection of relevant data. However, it is common that researchers refine the research question during their work (Voss et al., 2002).

Before conducting case study research, researchers need to specify which type of case study design to do. R. Yin (2009, p. 85) differentiates four types of case studies. The four types are differentiated by the number of case studies, and the number of units of analysis. Research focusing on only one case is called single-case design, in contrast to a multi-case design with more than one case study. Given the required resources and access to good cases, multiple-case designs are favored over single-cases. Conclusions independently originating from at least two cases are more reliable than those arising from a single case alone. When only two cases are planned, R. Yin (2009, p. 105) suggests selecting cases with very similar conditions (direct replication) or contrasting situations (polar cases).

Triangulation increases the validity of the findings of case study research. The opportunity to integrate data from different sources and research methods is considered a major strength of case study research. Triangulation may comprise methods such as questionnaires, observations, analysis of documents, historical data, and various forms of interviews (Eisenhardt, 1989; Voss et al., 2002).

5.1.2 Case Selection

Case selection was performed based on theoretical sampling. According to Eisenhardt and Graebner (2007, p. 27), theoretical sampling means that “cases are selected because they are particularly suitable for illuminating and extending relationships and logic among constructs.”

Following R. Yin (2009) and Eisenhardt (1989), who suggest favoring multiple-case designs over single-cases, this research has selected three companies for case study research in total. Additionally, two expert interviews with senior academics in the field of data utilization in manufacturing have been conducted. Both interview partners have extensive experience and a sound reputation in the field of data analytics in manufacturing and can contribute insights from different projects and companies. Therefore, the experts can serve as “a surrogate for a wider circle of players” (Bogner et al., 2009, pp. 1–2).

Like the three company cases in chapter 5.2, the expert interviews are presented and discussed separately, at first, in chapter 5.3. Afterward, the three case studies and the two expert interviews serve as input material for the cross-case study analysis in chapter 5.4.

The three companies have been selected due to the fact that they have been awarded as SP companies during two benchmarking studies. One of the studies was focusing on the future of lean and the other on digital technologies in manufacturing. The case selection is based on the assumption that these companies are better suited to learn from than randomly collected companies, and thus well suited for case study research (Eisenhardt, 1989, p. 537).

The selection process was the same for both benchmarking studies. In the first step, the results of an online survey were evaluated. Based on consistent and predefined criteria (e.g., level of lean deployment, kind, and maturity of digital technologies applied), researchers have identified 10–12 SP candidates. Afterward, semi-structured interviews were conducted with the candidates to validate the answers given in the online survey and to gather complementary and in-depth information. An anonymized case study was prepared for each SP candidate. In the second step of the selection process, industry professionals from the field of lean, operational excellence, and digital technologies selected up to five companies, which they consider to be mature companies other companies can learn from.

This two-stage process ensures that on one hand, all SP companies fulfill consistent and predefined criteria, and on the other hand that they are considered as role models also from a practitioner's perspective too.

By selecting the cases from the sample of SP companies of these two benchmarking studies, particular expertise in both fields, lean and digital technology, is ensured.

Company C is not a manufacturing company but is in the ICT business. Therefore, it does not fit the definition of DBAs as applicable to manufacturing, however, the company has included it mainly for two reasons. First, as an ICT company, using digital technologies to collect and process data is nothing new, as it's been daily business for many years. Consequently, the company is likely to have a high level of maturity of data utilization, which may be superior to those of manufacturing companies. Second, it is interesting to contrast the challenges and enablers of data utilization identified in an ICT company to those identified in the manufacturing industry in order to find commonalities and differences.

The three companies are presented in anonymized form in chapter 5.2.1. Since all companies employee at least 5,000 employees, all companies are large companies, thus constraining the variation of results due to the size of the case company (Eisenhardt, 1989, p. 537).

There is no clear consensus about the minimal, maximal, or right number of case studies in the literature. While Eisenhardt (1989, p. 545) suggests between four and ten case studies, Meredith (1998, p. 452) deems two to eight cases appropriate. R. Yin (2009, p. 105) agrees that already two cases might be enough. In general, for a limited set of resources, fewer case studies allow more depth of observation (Voss et al., 2002, p. 201). However, a small number of cases limits the opportunity for cross-case analysis and thus generalizability. Eisenhardt (1989, p. 533) presents the concept of theoretical saturation to determine the appropriate number of cases. Consequently, a sufficient number of cases is reached, when marginal additional insights from the next case becomes small. In this dissertation, reaching theoretical saturation for specific challenges was not possible during the three cases. However, a convergence of case study results regarding more general aspects, such as required employee qualification, was observed. To enhance the generalizability of the case study results, the findings are complemented with the results of two expert interviews in chapter 5.4.

5.1.3 Data Collection

Data triangulation—combining several data sources to study a phenomenon—improves the validity of findings of case study research (Eisenhardt, 1989, p. 538; Voss et al., 2002). For this research, an online survey, semi-structured interviews, workshops, and on-site visit observations were used as primary sources of data. Internal documents, such as presentations, reports, and publicly available information from the internet, have been used as a source of complementary information.

All companies selected for this case study research are SP companies from the two benchmarking studies outlined above. The respective online survey of the study serves two functions. First, the survey results already contribute to the data basis for the case study, and second, it serves as the foundation for the selection of SP candidates.

The survey questionnaire was developed in a joined exercise of industry representatives and academics, thus ensuring practical relevance as well as academic rigor. The initial version was tested with industrial companies, especially regarding the structure of the questionnaire, clarity of the questions, and duration. Feedback was integrated and the final version of the questionnaire was sent to more than 500 companies, mostly located in Central Europe. For the benchmarking study “Lean2020 – The Future of Operational Excellence,” focusing on the status quo as well as the future of lean, five SP companies have been selected from a total sample of 75 companies (see chapter 3.1). For the benchmarking study “Digital Technologies – Evolution of production in high-wage countries,” focusing on selected digital technologies such as MES and Big Data analytics, five SP companies have been selected from a total sample of 139 companies (Benninghaus, Elbe, Budde, & Friedli, 2018).

Semi-structured interviews have been the most important source of information of the case study research. After initial interviews as part of the benchmarking procedure, additional interviews have been conducted with a dedicated focus on the research at hand. Five dedicated interviews have been conducted, ranging from 70 to 130 minutes in duration. To structure the interview, an interview guideline was developed and distributed to the interviewee prior to the interview (see Appendix C). Following Saunders, Lewis, and Thornhill (2009, p. 320) semi-structured interviews are not standardized. Instead, the researcher has a selection of questions and topics to be covered. However, this selection can vary from interview to interview. Considering the specific context, some questions may be omitted while others are added. Also, depending on the flow of the conversion, the order of questions may be altered.

With the exception of one interview, all discussions have been audio-recorded, allowing the interviewer to focus more on the interview than on note-taking. Notes from the interview have subsequently been cross-checked with audio recordings and updated. The refined and detailed interview notes were sent to the interview partner for confirmation or clarification if necessary.

Workshop and site visit observation provides additional information and contributes to data triangulation. All SP companies have been visited at least once for a full-day on-site visit, thus allowing direct observation at the site. Moreover, as several representatives of the hosting company have been present, the risk of single response bias found in interviews is reduced. Finally, additional information from internal and publicly available documents have been integrated into the case studies. These documents comprise internal reports and company presentations as well as publicly available marketing publications and annual reports.

5.2 Within-case Study Analysis

This research follows the case study approach described by Eisenhardt (1989). Analyzing the data consist of two parts: within-case analysis and cross-case analysis. The objective of the within-case analysis is to gain familiarity with the collected data of each case as a standalone entity. Within-case study is also motivated by the risk of “death by data asphyxiation” (Pettigrew, 1990, p. 281) due to too much data. A within-case study consists of detailed descriptions of each case and supports researchers to deal with the large volume of data in the early phase of the data analysis. Starting with a focus on cases as a standalone entity allows for identifying patterns that are unique to the case before searching for generalizable patterns across several cases. Also, by gaining familiarity with the data during within-case analysis, cross-case comparison can be conducted faster (Eisenhardt, 1989, p. 540).

5.2.1 Overview

Table 20 presents an overview of the three case study companies.

Table 20: Case overview

	A	**B**	**C**
Company	**Automotive Supplier Company**	**Automotive Supplier Company**	**Information and Communication Technology Company**
Market	B2B & B2C	B2C	B2B & B2C
Employees	> 100,000	> 5,000	> 20,000
Revenue in USD	> 15 bn	> 2 bn	> 10 bn
Scope of case	Plant	Company	Company

The three companies present several commonalities. First, all three companies have been selected as SP companies. Second, all companies are (or among) leaders in their respective markets. Third, they all serve the B2B market, while two also operate in the B2C market. Fourth, with a minimum number of employees above 5,000 and an annual revenue above $2 billion, all companies are large companies, operating and selling their products and services internationally. Fifth, all companies can be described as technology companies. Sixth, all companies are located in Central Europe and also generate a major share of their revenues in Europe. Company A and

Company B are both automotive supplier companies, but develop and produce different products.

Anonymity was granted as a precondition for case study interviews. Thus a more detailed description of products must be dispensed with. Having the high pressure on cost efficiency in the automotive supplier industry and the fact that lean and data utilization both aim at reducing waste and increase efficiency in mind, it is not surprising that two SP companies are from the automotive supplier industry.

On the other hand, there are also some differentiating characteristics. Whereas Company B and Company C are portrayed from a company perspective, the scope of Company A is the manufacturing plant. Furthermore, whereas the main interview partners of Company B and Company C are working in corporate functions, both main interview partners from Company A have a site role.

Company C is a special case, as it does not operate a manufacturing system in the classical sense of converting physical raw materials to finished products. Instead, Company C offers ICT services to different industries as well as private end customers. Therefore, some aspects of the questions might not be equally applicable to Company C. However, challenges of data utilization regarding employee and organizational enablers may be similar to those of manufacturing companies and the comparison might yield interesting commonalities or differences. Hence, Company C can be considered to a certain degree as a polar type case (Eisenhardt, 1989, p. 537).

The cases vary in length and depth, depending on the level of interaction (e.g., joint projects, workshops, site visits, case interviews) or other forms of exchange as well as on the relevance for the research at hand. All cases follow the same basic structure, however, the specific content may vary due to different contexts. The case description comprises the following sections: general information (1), lean status quo (2), strategic alignment of lean and digitalization (3), use of data-based applications (4),[23] data utilization use case (5), challenges and enablers of data utilization (6), impact of data utilization on lean (7), and key implications (8).

The section *general information* provides background information on the industry the company is working in, its size in terms of the number of employees and the annual turnover in 2018. Also, the role of our main contact person is outlined as well as specific peculiarities discussed. The section *lean status quo* evaluates whether the company is organizing its operations according to lean principles and which of the 10 key lean practices is actually applied in the company. This information is relevant to assess the ability of the interview partner to evaluate the impact of DBAs on the lean practices.

[23] Only in case I and case II

Lean and digitalization[24] pursue similar goals, including high transparency, waste reduction, and customer value creation. Some publications claim that the integrated consideration of lean and digitalization in manufacturing yields considerably higher potentials for efficiency gains than if one of the approaches is implemented alone (Küpper, Heidemann, Ströhle, Spindelndreier, & Knizek, 2017, p. 2). This observation motivates the question if companies already align both concepts. Thus in the section *strategic alignment of lean and digitalization*, the interviewees are asked whether their company has an integrated strategy and whether lean and digitalization responsibilities are organizationally integrated.

Section 4, *use of data-based applications,* addresses the question of which of the identified DBAs in chapter 4.1 are actually used in the industry. The section *data utilization use case* presents a self-selected, recent use case of data utilization. Based on this specific use case, the following questions were discussed. First, what are the main drivers that motivated the use case? Second, what are the objectives? And third, the interview partners are asked to outline the most critical challenges and enablers from the following three perspectives: technological, organizational, and employee qualification (Hirsch-Kreinsen et al., 2018, p. 181). Also, to evaluate the validity of the TAM theory introduced in chapter 1.5.5, acceptance problems among employees and potential mitigation strategies are discussed.

The scope of the section *challenges and enablers of data utilization* in manufacturing is very similar to the section before, but instead of focusing on a specific use case, the scope is broader. This section discusses challenges and enablers, as well as acceptance issues and mitigation strategies from a general perspective. The final section, *impact of data utilization on lean*, was designed to discuss DBA – lean practice combinations, where a high impact is assumed. The feedback has been integrated into the DBA – lean practice impact evaluation in chapter 4.2.

Finally, the last section summarizes the key implications from each case as the basis for the cross-case study analysis in chapter 5.4.

5.2.2 Case I: Company A

5.2.2.1 General Information

Company A is a global technology company with several divisions. Both interview partners of this case have answered the interview questions from a site perspective. The site belongs to a division producing components for powertrains in the automotive industry. Therefore, Company A is described in the overview and in the following as an automotive supplier company. The figures on the number of employees and

[24] The term digitalization refers in this context to the use of modern, digital technologies, such as the core technologies presented in chapter 2.3.3.

revenue, however, reflect the figures of the whole company. In total, Company A employs more than 100,000 employees worldwide and has generated an annual turnover in 2018 of over $15 billion. The site represented in the case study is one of the most mature sites regarding the integration of SM technologies within the manufacturing network, which comprises more than 50 sites worldwide.

For the case study, specific interviews have been conducted with two representatives of Company A in addition to the insights from the benchmarking. Representative 1 is a key employee for integrating industries 4.0 solutions and data analytics into the existing production system. Representative 2 has a focus on internal material flow and logistics and is responsible for piloting connected logistics solutions and the digitalization of value streams. The interviews have been conducted separately, but all data gathered are combined in this case study.

5.2.2.2 Lean Status Quo

Company A operates a company-specific production system that follows the principles of the TPS. Lean, however, is not only considered as an approach to minimize wasteful activities but is also described as the basis for innovation. Clear standards are essential prerequisites for connected manufacturing and connected logistics. The five lean principles introduced in chapter 2 are guiding principles of the company's LPS. To ensure continuous improvement, Company A applies system-CIP (continuous improvement process) as a method for the process and value stream design (at least four times per year) and point-CIP for process stabilization and improvement (every week).

Of the 10 lean practices, serving as the foundation for the DBA-lean impact evaluation, all 10 are applied at the site at different implementation levels. Preventive Maintenance is used, including first applications of predictive maintenance. For Quality Management DataMatrix-Code track-and-trace data are collected for most but not all products. Continuous Flow Production is an objective of Company A, but at the site partially challenging to realize due to building restrictions. Therefore, the site seeks to implement smart intralogistics solutions to enable flow. The Pull system is used whenever applicable. Quick Changeover Techniques are applied as well as measures to reduce the lot size. Value Stream Mapping is performed at least once a year and always in case of process changes. It is supported by a dedicated LPS unit. CI and employee involvement is fostered by an idea suggestion system, support for creating CI suggestions, and a small financial incentive, even if the idea is not implemented. Cross-functional Teams are formed for time-critical projects and rollouts. Self-directed Work Teams are not very common, but there are some scrum teams working in a self-directed mode.

5.2.2.3 Strategic Alignment of Lean and Digitalization

Company A has both a strategy to implement lean as part of the company's LPS and a formalized digitalization strategy. The digitalization strategy is formulated at the highest level of the division. The strategy provides a framework for production sites. Depending on the current need and maturity, elements of the strategy are prioritized over others. Striving for the highest level of digitalization is not always aspirational. The main objectives of the digitalization strategy are the digitalization of value streams, the implementation of AGV systems, and a high level of MES connection of the production equipment for automated data collection.

The value proposition of these efforts is increased transparency, increased equipment availability, and the integration of production and logistics. Data collection and integration is expected to yield new insights that have been hidden hitherto. On the site level, the site leader is formally responsible for digitalization initiatives but has delegated the responsibility for the operative design and implementation of digitalization initiatives to interview partner 1.

From an organizational embedding perspective, lean and digitalization are separated. Lean know-how is concentrated at a central unit for the company LPS and digitalization expertise is concentrated at a unit dedicated to the MES and data analytics system. Also, an integrated strategy, taking lean and digitalization into consideration simultaneously, does not exist. However, although both units are separated organizationally, they often interact closely in projects. Within Company A, lean is considered as the basis for digitalization. Therefore, when introducing new production equipment, the lean unit defines the process to ensure high OEE and low inventory levels. The lean process is then digitalized by mirroring the process in the MES system.

5.2.2.4 Use of Data-based Applications

From 12 DBAs[25] identified in chapter 4.1, Company A has already eight DBAs in use, and three more are currently in the piloting and testing phase (Table 21).

Table 21: Overview: use of data-based applications – Case I

Overview: Use of Data-based Applications			
Layout Planning	*Production Scheduling*	Real-time Control	System Performance Measurement
Condition Monitoring	Predictive Maintenance	*Prescriptive Maintenance*	Track and Trace
Material Flow Management	*Inventory Management*	Product Quality Monitoring	Product Quality Improvement
Key	Not in use	*Testing phase*	In use

Real-time Control is used to enable a faster reaction to problems. In the event of a disturbance, employees receive relevant information on the problem and its location directly on a mobile device. The availability of spare parts can be checked and the delivery issued remotely. *Real-time Control* has contributed to a reduction in downtime of 20 percent.

System Performance Measurement is applied as part of a digital shop floor management. Manufacturing data are tracked and visualized to increase the transparency of the value stream. Consistent definitions of KPIs across several manufacturing sites allow comparisons across the manufacturing network. To minimize human effort, data collection and processing are already automated to a high degree, with the intention to further decrease the share of human collected data. Information collected and visualized by the application *System Performance Measurement* is accessible remotely to everybody with access rights.

Track and Trace is used as well as smart *Material Flow Management* systems; for instance, e-Kanban and AGVs with automated generated path planning. *Condition Monitoring* and *Predictive Maintenance* is implemented for some machines. For example, the vibration of a spindle is monitored to detect patterns and derive preventive maintenance plans. *Prescriptive Maintenance* is not in place. *Product Quality Monitoring* is supported by collaborative robots, which perform visual product quality inspection. Data collected by the MES system is accessible to the quality department. By using product and process data, root-cause analysis is supported.

[25] At the time of the interviews, the two DBAs of the EHS category were not yet part of the collection.

5.2.2.5 Data Utilization Use Case: Manufacturing Analytics Solution

During the case study interview, one or two use cases of data utilization were discussed in detail. The interview partners were free to select the use case and were not restricted to the DBAs identified in the literature. The use cases were selected based on two criteria. First, the relevance of the use case for the company and second, the personal involvement of the partners in the project, which ensures a sound knowledge base on the objectives, challenges, and enablers of the use case.

Company A presents a use case of the *System Performance Measurement* application. Initially, the process of data integration from different sources was extremely tedious and time-consuming. Employees needed to spend several days, in an extreme case, up to 50 days, for locating and merging data from different sources and databases. Many ideas for data analytics were not pursued further, as the initial effort of data collection was too high. As a result, the plant initiated the project *Manufacturing Analytics Solution* (MAS) to enable easy and fast access to distributed data via a central platform. The proposition that the MAS would decrease the effort of data analytics applications significantly convinced the management to invest in the technological infrastructure.

The main components of the technological infrastructure were scalable data storage as well as powerful and scalable computing power. The MAS system was realized with a distributed Hadoop system. Only a few new sensors needed to be installed, as the majority of the relevant manufacturing data was already collected.

Major challenges for the MAS project were to create full transparency on the data integrated into the system, to ensure data quality and data integrity, to ensure employee acceptance of the new system, and enable employees to use it. Data-integrity includes the standardization of metrics and failure codes across several sites to enable consistent data architecture across the production network. This standardization must be driven by management, as there is little intrinsic motivation of the sites to do this exercise.

Having the technological infrastructure and data architecture in place, a critical enabler of data analytics application was to bring together manufacturing data from the MAS and manufacturing domain knowledge of employees for data interpretation. To empower employees to use the system, basic training was provided to employees by internal MAS experts, so-called Citizen Data Scientists.

Citizen Data Scientists are plant employees without formal education as data scientists, but a strong intrinsic motivation to build necessary skills for data analytics. They often have a high IT affinity and some IT skills they have gained in a private context. Citizen Data Scientists receive extensive on-the-job training in data analytics tools. In addition, they may reduce their amount of daily workload to have time for individual and self-organized training, including the internet and YouTube as a source of training material.

Besides the basic training, employees are supposed to familiarize themselves with the opportunities of data analytics in workshops aiming at interactive learning by hands-on testing of their own ideas. Afterward, the Citizen Data Scientists serve as sparring partners and provide support to employees who propose a data analytics approach to solve a problem within their working area.

From an acceptance perspective, most employees welcome the new MAS system, especially younger colleagues who desire to improve the previous situation of tedious data integration and the work with local data in excel. Interestingly, most persuasions were needed to convince employees of the value of the new system, that have worked a lot with data in the past and are considered as experts in the field of data analytics. The acceptance of the MAS system by employees depends to a high degree if the individual perceives the analytic solution as an improvement for his or her daily job or as an act of paternalism, decreasing his or her autonomy.

A second driver of acceptance is active involvement in the development of the solution. For instance, a maintenance employee is much more likely to accept a data analytics solution for preventive maintenance if he or she was involved in the development of the solution. On the contrary, the willingness to follow data analytics results instead of personal know-how to derive maintenance actions decreases, if the solution was developed without personal involvement and if it was introduced top-down by the management.

Although not every employee can be involved in the solution development process equally, our interview partner stresses the point that each employee must have at least the opportunity to provide feedback: “Introducing a new solution without collecting feedback first does not work.” In Company A, workshops are conducted if new solutions are integrated into the process and these workshops usually comprise a basic introduction to the solution and the opportunity to provide feedback. Thereby, all affected employees can provide concerns and improvement suggestions.

A general barrier for data utilization in manufacturing is the potential to use the data for individual performance measurement or even individual behavior control. Necessary to overcome this barrier is early and open communication with the employee representation, to create transparency about which data are used for what purpose and to define clear limits for tracking individual behavior and performance. Without approval of the employee representation body, specific DBAs may not be implemented.

The maturity of the MAS project is currently at approximately 70–80 percent. After a successful rollout in the plant of this case, it is intended to also be implemented in other sites of the division.

5.2.2.6 Challenges and Enablers of Data Utilization

The last section has presented challenges and enablers of data utilization in manufacturing, linked to the presented use case of MAS. This section discusses the question of challenges and enablers of data utilization more broadly and detached from the specific use case. The interview partners were asked to outline challenges and enablers from three categories: technology, organization, and employees. The focus, however, is chosen by the interviewee, depending on the individual perception of relative relevance and criticality. As some challenges and enablers are mentioned in the discussion of the data utilization use case and as general challenges and enablers of DBAs, occasional duplications occur.

Technological Challenges and Enablers

Both interview partners put little emphasis on technological challenges and enablers. It is common sense that the collection, transfer, and storage of data requires an appropriate IT system that is reliable and able to process large amounts of data quickly. Therefore, Company A uses a scalable Hadoop system that allows the carrying out of intensive computing processes with large amounts of data on distributed computer clusters. Furthermore, data security is seen as an obvious challenge. Therefore, Company A not only invests in technological protection measures against external hackers but also in training for employees to avoid data or access credential theft by social engineering or phishing. Access and identification security are part of mandatory training for all employees regarding data security. To ensure data integrity, Company A strives to automatically collect as many data points as possible with an MES system. Currently, this applies to more than 80 percent of the data. MES data meet the VDA[26] requirements of data integrity and therefore are considered as highly reliable.

Besides meeting these basic technological requirements, Company A enables data analytics by providing a smart analytics platform, which is designed and implemented by central units in collaboration with the manufacturing experts in the plant. As seen in the example above, this platform provides a powerful basis to perform targeted analysis without spending days for data collection and individual solutions programming in excel. The reduced effort substantially increases the motivation of manufacturing employees to think about data analytics opportunities and build a prototype to test the idea.

Organizational Challenges and Enablers

According to the interview partners, enabling sustainable data utilization projects requires a systematic process to identify promising opportunities. Therefore, in the

[26] German Association of the Automotive Industry

respective plant, *Analytic Business Case Reviews* are conducted. During these reviews, the responsible for digitalization in the plant, who is also the head of the internal Citizen Data Scientist expert network, presents tested and rolled-out solutions from other areas to a value stream manager. He thereby acts as a competent advisor as well as a sparring partner.

If the value stream manager recognizes enough potential for similar applications in his or her area of responsibility, potential analytic projects and support of the internal Citizen Data Scientist network are discussed. Given the intention to implement a selected analytic project, objectives and resources, including financial and human effort, of the project are defined. Only if enough resources can be dedicated to the project, the implementation starts. If necessary, a higher-level committee, comprising the site leadership team, can be asked to provide additional resources.

The approach of setting an Analytic Business Case Review as a starting point of an analytic project has several advantages. First, the expertise of data analytic projects is combined with expert knowledge on the value stream at hand. It will only start if the digitalization responsible and the value stream manager jointly recognize enough potential in a project.

Second, by having supported several data analytics projects before, interview partner 1 has a good feeling for the required human and financial effort, thus supporting a realistic cost-benefit consideration. Also, this experience ensures that projects are staffed adequately to achieve the project objectives in time.

Third, data analytic projects are implemented only if the responsible value stream manager requests the implementation. The projects are primarily implemented by workers of the respective value stream, supported by the Citizen Data Scientist network. By keeping the decision and implementation on the operative level, the acceptance of the solution among shop floor employees is increased.

Another positive aspect is that the Analytic Business Case Reviews link different management levels (operative shop floor management, staff unit for digitalization, and the site leadership team), while the internal Citizen Data Scientist expert network creates links on the operative level.

Company representative 2 adds further organizational requirements. He argues that due to the increasing use of manufacturing data, more decisions are taken automatically. For instance, AGVs already optimize their route planning autonomously. Nevertheless, the company should not entirely rely on the robustness of the system but rather prepare itself for the case of disturbances. Therefore, he asks for a risk assessment and an emergency concept, so that employees are prepared to react quickly and efficiently in case the data-based self-optimized systems crashes.

To ensure acceptance of employees and avoid legal conflicts, Company A has established clear standards, detailing which data is allowed to be tracked and which is not allowed to be tracked. While employees are welcoming digital technologies and data analytics in general, acceptance is limited if data can be used to monitor individual behavior and performance. To be compliant with data security and privacy regulations, each plant has its own Data Security Officer.

Employee Challenges and Enablers

Providing a basic understanding of DBAs to shop floor employees is key for the ability to use the application effectively and to ensure acceptance of the application among the prospective users. What is new and not understood may create uncertainty and rejection. Also, only by understanding the logic behind a data analytics solution, shop floor employees can provide qualified feedback to the solution developer. Consequently, all involved employees are trained when a new data analytics solution is introduced.

Usually, an existing data analytics solution needs to be adapted to the specific process. End-users are encouraged to participate in the adaption process to bring in firsthand process knowledge. Thereby, the end-user is also part of the solution development and as such, shows a higher interest in the successful implementation of the solution compared to the implementation of a top-down provided, ready-to-use solution. Basic training is essential to enable shop floor employees to be part of the application development, which in turn contributes to the acceptance of the result.

The critical importance of the acceptance of a new solution is illustrated by a preventive maintenance example. A new data analytics solution was presented to support maintenance employees in predictive maintenance. However, experienced maintenance employees rejected the proposed solution and argued that they preferred to rely on their longtime experience instead of on suggestions from the new tool. Although the analytic solution worked from a technological perspective, it was never successfully integrated into the existing maintenance process.

A key takeaway for the management was to involve all stakeholders early in the development process of data analytics solutions. The solution developer and all relevant stakeholders on the job floor need to discuss and agree which action needs to be triggered if the predictive maintenance application detects a particular signal. Taking this lesson into consideration, another DBA use case of quality monitoring was jointly developed and successfully integrated into the process. Representative 1 summarized it as follows: “The main challenge of an analytic solution is often not technical, but a question of employee acceptance.”

In terms of employee qualification, it is neither feasible nor necessary to provide advanced training in data analytics for every employee. As discussed before, Company A instead selects few especially motivated employees with preexisting basic knowledge of IT, programming, and data analytics and provides extensive training to those individuals to become Citizen Data Scientists. Citizen Data Scientists are internally referred to as the *masters of the data.* Unrelated to a specific use case of data analytics to be implemented, there is currently no standard training in data analytics for all employees. However, the topic of data utilization in manufacturing, especially data security, is addressed to some degree in compulsory training on industry 4.0.

The training of the Citizen Data Scientists rests on three pillars. First, they receive basic training comprising different training modules provided by the company. Second, as new data analytic techniques and approaches are developed frequently, self-learning is a central element. Citizen Data Scientists are encouraged to search and use non-standardized learning material, also including, for instance, internet blogs and YouTube tutorials. To be effective, self-learning must be supported by the supervisor (e.g., by providing additional resources to the employee, such as time dedicated to training and, if required, a budget for fee-based training). The third pillar comprises the ongoing exchange of Citizen Data Scientists with experts of Company A's central units (e.g., professional data scientists). Within the plant, knowledge and experience is supposed to be shared during regular meetings of the Citizen Data Scientists network.

Citizen Data Scientists combine manufacturing domain know-how with data analytic skills. This combination is highly appreciated by external companies but also central units of Company A. As a result, Citizen Data Scientists receive an above-average number of job offers.

Hiring professional data scientists from the job market is challenging due to intense competition for few available experts. In comparison to banks and consultancies, a manufacturing company often struggles to offer competitive salaries. However, as one interview partner proudly points out, the manufacturing site allows testing new solutions in a physical setting within 100 meters of the working place. The opportunity to easily test an idea and receive feedback instantly is seen by many candidates as a strong argument for Company A, although higher salaries may be offered elsewhere. Also, Company A has a good reputation as a reliable employer, a fact that also facilitates the hiring of highly demanded data scientists.

5.2.2.7 Impact of Data Utilization on Lean

“Lean remains lean” in terms of the basic philosophy of waste reduction. Company A does not expect major changes of lean due to DBAs on the level of lean principles, such as pull or flow. However, some implications are expected on the level of lean practices. Interview partner 1 suggests increased transparency, preventive maintenance, and internal logistics as examples of lean being supported by DBAs.

Regarding transparency, data can be used to evaluate the result of performed improvement measures to a machine. Prior to the introduction of the MAS system described above, the effort for this use case was too high, with the result that there was some vagueness whether or not the taken improvement measures have been useful. Furthermore, increased transparency is expected to identify unnecessarily high inventories, which then can be optimized. A current project is concerned with the simulation of material flows, integrating real-time inventory data.

After initial barriers, the DBA *Predictive Maintenance* (monitoring the vibration of a spindle to detect patterns) has been implemented successfully, thus supporting the lean practice Preventive Maintenance. In regard to internal logistics, Company A uses milk runs as one element of internal material distribution. Real-time data from the production system allows changing the takt from static to dynamic, based on actual demand. Thereby, the objective is to reduce the number of milk run tours, while still ensuring timely delivery of materials to the machines.

DBA – Lean Practice Impact Evaluation – Industry Feedback

Interview partner 2, serving as an specialist for internal material flows, was asked to evaluate three DBA – lean practice combinations with a theoretically high impact. The evaluations given in the interview have been integrated into the impact evaluation of DBAs on lean practices in chapter 4.

The first combination discussed was the DBA *Production Planning*, and the lean practice Pull/Kanban. From a theoretical perspective, a contradiction between push, resulting from central production and material flow planning, and demand-oriented pull can be derived (see chapter 4.2.3.3.2). And in fact, in the plant of Company A, the pull system is substituted at least partially as the production of critical parts is planned centrally instead of being triggered by a Kanban card. However, apart from the production of critical parts, the material flow and replenishment of supermarkets is managed by a pull system. Central planning and the Pull/Kanban system are considered more as complementary elements of the systems than as opponents.

The second theoretical proposition discussed was that by having access to real-time data, such as current demand and material availability, the DBA *Material Flow Management* could strongly support the lean practice Pull/Kanban. This proposition was supported by the logistics specialist.

The Pull/Kanban is realized in Company A as e-Kanban. Data on consumed material and projected demand are feedbacked from machines and current material stock feedbacked by smart supermarkets to the ERP inventory management system. Based on these data, the system calculates the optimal lot size for replenishment. Thus, the e-Kanban is controlled by a virtual representation of the physical material flow. Major advantages over traditional Kanban systems are a fast and secure transfer of Kanban signals as well as the ability for dynamic lot size adaptions.

Thirdly, interview partner 2 was asked to evaluate the potential of the application *Track and Trace* to support the lean practice Value Stream Mapping. Literature suggests that *Track and Trace* and other manufacturing data can be used to perform a cost-efficient assessment of the current value stream, an approach referred to as "VSM 4.0" (see chapter 4.2.3.3.5).

The concept was very well received by the interviewee. He especially appreciates the perspective to perform several value stream mappings for different product variants with reduced human effort. Furthermore, due to a wide variety of products, bottlenecks may not be static but depending on the product. Such dynamic changes are likely to be better reflected in repeated and data-enriched value stream maps compared to standard value stream maps. However, VSM requires a physical presence on the shop floor. Merely relying on manufacturing data without being at the place of the process significantly increases the risk of data misinterpretation. In summary, a virtual VSM is rejected, while data supported traditional VSM is perceived as a worthwhile approach to enhance the current VSM methodology.

5.2.2.8 Implications Case I : Company A.

Table 22 presents key implications of the first case study. They serve as input for the cross-case analysis in chapter 0. Findings from the discussion on DBA - lean practice impacts have been included in chapter 4.2 and are not listed in the table below

Table 22: Key implications – Case I

I	Company A - Automotive Supplier Company	P
I.1	Company A has a lean strategy as part of the LPS and a formalized digitalization strategy on a division level. Lean is a foundation for digitalization.	123
I.2	Lean and digitalization are not combined in a common strategy and are separated from an organizational embedding perspective. Lean and digitalization experts mainly work together in projects (e.g., when introducing new production equipment).	123
I.3	Company A already uses eight out of twelve discussed DBAs at least partially in the case plant. Three more DBAs are currently in the testing phase.	124
I.4	The main barrier for data analytics application was an immense effort to find, combine, and analyze relevant manufacturing data. This barrier has been overcome by implementing the Manufacturing Analytics Solution (MAS) system.	125
I.5	Key technological requirements of the MAS are scalable computing and storage capacity. The MAS system is built on a distributed Hadoop system.	125
I.6	A key challenge for the MAS was to ensure data quality and data integrity by setting plant or even company-wide standards. Enforcing the standardization is a management task.	125
I.7	The concept of Citizen Data Scientists is useful to build integrated manufacturing and data analytics expertise internally. Training of Citizen Data Scientists includes regular training, self-training, and regular exchange with plant and central unit data analytics experts.	125
I.8.	Active involvement in the development is a driver of the end-user acceptance of analytic solutions. “Introducing a new solution without collecting feedback first does not work.”	126
I.9	A general barrier for data utilization in manufacturing is the potential to misuse data for individual performance measurement and individual behavior control.	126
I.10	The identification and implementation of DBAs should follow a structured approach, such as the Analytic Business Case Review approach.	127
I.11	To create transparency on data utilization, Company A has developed clear data utilization guidelines. To ensure compliance with internal and external data protection requirements, every plant has a dedicated Data Security Officer.	129
I.12	Training for DBA end-user, including the basic logic behind the DBA, is essential for three reasons: to enable effective use of the application, to enhance end-user acceptance, and to facilitate qualified feedback to developers. Training is provided to the relevant stakeholders if a new solution or new equipment is introduced.	129

I: Implications, P: Page

5.2.3 Case II: Company B

5.2.3.1 General Information

Company B is a major automotive supplier company, headquartered in central Europe. It has more than 5,000 employees worldwide and has generated an annual turnover in 2018 of over $2 billion. Since 1980, the turnover has grown significantly through internal growth and external acquisitions with a compound annual growth rate of 13 percent. The company operates more than 15 R&D and production facilities in the EU, Asia, North and South America. The main contact is Head of Lean Management. He has answered the interview questions from a corporate perspective.

5.2.3.2 Lean Status Quo

Lean is considered a tool to ensure a basic level of standardization among the sites of the production network, however, lean is also seen as a philosophy to foster continuous improvement and employee involvement in CI activities. Current objectives of lean management in Company B are quality improvements by failure reduction, thus reducing failure costs and improving delivery capability.

Company B operates a company-specific production system that is based on lean principles. The five key objectives are safety, quality, delivery, cost, and sustainability. The company-specific LPS visualization has the PDCA circle at its core, thus highlighting the importance of CI within the company. CI is part of the daily job and all employees are expected to be actively involved in CI activities, especially during the implementation phase of new processes. During a site visit, a manager revealed that at the site, on average, two improvement ideas are contributed for every employee, which amounts to 2,000 improvement suggestions per year in one plant.

From the lean practices described in chapter 2.2.3, all 10 are used at Company B to varying degrees. All 10 lean practices are well-known to the interview partner, so he is considered to be qualified to provide industry feedback in the DBA – lean practice impact evaluation.

5.2.3.3 Strategic Alignment of Lean and Digitalization

Company B has a digitalization strategy with a strong focus on a cross-plant MES system with comprehensive analytics functionalities. Implementing a company-wide MES system needs to be managed by corporate, to ensure standards and provide sufficient resources. On a site level, individual digitalization projects are developed, tested, and implemented, but on a comparably small scale.

Regarding the integrated consideration of digitalization and lean, our interview partner highlights the importance of close interaction between IT teams and the lean team. The ideal process from an IT perspective does not necessarily equal the perfect process from a practitioner's perspective on the shop floor. As a result, the lean unit is needed to ensure the applicability and user-friendliness of IT tools for the company LPS. Nevertheless, the digitalization strategy is currently not aligned with the lean strategy embodied by the LPS.

Also, from an organizational embedding perspective, the lean responsible and the responsible person for the MES system are separated. The interviewee argues that the integrated consideration of lean and digitalization is the responsibility of the top management. This responsibility is often delegated to project teams, for instance for the implementation of a new MES system. The MES project team was supported by the corporate lean unit, thus ensuring to have both perspectives in the team.

5.2.3.4 Use of Data-based Applications

Company B has eight out of twelve DBAs discussed in the interview already in place, with two more DBAs in the testing phase. The overview shown in Table 23 applies to Company B as a company, whereas the individual sites may have fewer DBAs in use.

Table 23: Overview: use of data-based applications – Case II

Overview: Use of Data-based Applications			
Layout Planning	*Production Scheduling*	Real-time Control	System Performance Measurement
Condition Monitoring	Predictive Maintenance	Prescriptive Maintenance	Track and Trace
Material Flow Management	*Inventory Management*	Product Quality Monitoring	Product Quality Improvement
Key	Not in use	*Testing phase*	In use

Data-based *Production Planning* was rolled out only recently with the primary objective of integrated planning across several sites within the production network. In the production network, site X produces parts used in site Y. Planning across sites allows increasing flexibility while reducing inventories at site Y. Currently, three production sites are linked to the central planning system.

For *Real-time Control* a central cockpit was designed that visualizes real-time data from process control sensors. Remote access to the dashboard is possible, given the access rights. As an example of the application *System Performance Measurement,* the KPI OEE of almost all production machines is measured and visualized on a remotely accessible dashboard.

Predictive Maintenance is currently tested at a few machines. However, currently, the effort and costs for *Condition Monitoring* and *Predictive Maintenance* are often exceeding the costs for preventive maintenance on a fixed schedule basis. Hence, those DBAs are not prioritized.

Track and Trace, in contrast, is widely rolled-out as traceability is a basic requirement of many OEM customers. *Material Flow Management* is realized by e-kanban and AGVs. In one plant, 25 AGVs are used to automate the intralogistics and since their introduction, significant fewer accidents have occurred and as a result, fewer goods were damaged.

As part of the DBA *Product Quality Monitoring,* critical to quality product characteristics are monitored and documented automatically. This documentation is requested by compliance requirements. Using product and manufacturing data for *Product Quality Improvement* is applied mainly in mechatronics manufacturing. Previously, the integration of several data has been very time consuming, which was a barrier to data-based quality improvement. Accessing and combining different sources of data is strongly facilitated by the new MES system.

5.2.3.5 Data Utilization Use Case: Global Performance Cockpit

The main objective of the *Global Performance Cockpit*, which is part of a new MES system, is to consolidate machine data to ensure delivery capabilities. To keep service level agreements, new equipment often needs to be fully utilized. When falling below an OEE threshold of 75 percent for some machines, the site's delivery capability is at risk. This, in turn, may have negative impacts on another site in the network, depending on the timely delivery of parts. Based on real-time data (updated every five seconds), the Global Performance Cockpit highlights critically low OEE performance levels early. The small latency time allows rescheduling in other production sites in time before the production is impacted by missing supply of intermediate parts.

In addition, the Global Performance Cockpit allows internal benchmarking between similar production equipment to detect problems, with low effort and even across sites.

The main technological requirement is to build the physical infrastructure for the system, including the installation of sensors at the production equipment and the definition of interfaces for data transfer to the central MES system. To make performance metrics comparable, they need to be standardized company-wide. This standardization exercise is an organizational challenge, as different sites may have used different KPI definitions historically.

A critical enabler is to train employees to understand the value of extra effort for data collection and the importance of providing data without errors to the system. Although most of the data is collected automatically, extra effort may arise for some employees for manual data input.

Therefore, change management is necessary to “take the people on board.” It is a leadership task to communicate the benefits of the system and explain why it is beneficial for the company (e.g., in terms of competitiveness) as well as for every employee.

For the successful implementation of the use case, ongoing leadership commitment was critical as large-scale IT projects as the implementation of an MES system are often complex, very expensive as well as time and resource consuming.

5.2.3.6 Challenges and Enablers of Data Utilization

Technological Challenges and Enablers

Out interview partner has identified three areas that might constitute barriers to implementing DBAs. First, enabling access to distributed data requires the integration of several data sources and IT systems. This integration is expensive and time-consuming. Second, IT system stability is fundamental for some DBAs. For example, production scheduling and control requires 100 percent IT system availability; otherwise, the whole production system stops running. Third, data protection is considered critical as hacking manufacturing data systems may allow competitors to gain business-critical information.

Organizational Challenges and Enablers

Organizational challenges and enablers comprise management commitment for DBA projects, effective change management, and the protection of personal data.

As discussed in the previous section, providing the IT infrastructure for DBAs in a large company is a highly expensive and effortful project, which can only be successfully completed with ongoing management support.

Moreover, management has a critical role in the change management process as some DBAs bear the risk to reduce the scope of a job description or even to make the job redundant. For instance, the Predictive Maintenance application can decrease the total amount of maintenance efforts and, thus, reduce the number of maintenance jobs. Fear of job losses is a strong driver for employee resistance. Consequently, management needs to communicate the value and needs of the application and offer employees a reliable perspective within the company in case of job reductions.

Data protection has not only a technological perspective but is also an organizational challenge. Management must ensure the protection of personal related data. Apart from the fact that companies are bound by law to adhere to personal data protection regulations, at least in the EU labor unions have the power to block the rollout of manufacturing data systems that enable monitoring individual behavior.

Employee Challenges and Enablers

The major importance of employee qualification to work with a new system or application has been discussed in the use case section. Employees need to understand the implications of the data they enter into the system. Our interview partner emphasizes that knowing the “why” of data collection significantly increases the quality of the manually entered data.

On the middle management level, a holistic understanding of IT systems and manufacturing systems is required to drive DBA projects. However, nobody can fully gain a complete understanding of both the IT system and the manufacturing system, thus bringing together IT and manufacturing experts in cross-functional project teams remains a critical enabler of DBAs.

5.2.3.7 Impact of Data Utilization on Lean

“Full transparency of data and processes on the shop floor with digital interfaces on the shop floor.” This statement summarizes the expectations of the integration of digital technologies in Company B's LPS. Full access to data will enable near real-time analysis capabilities and thus supporting the lean continuous improvement process. An example provided by Company B are six sigma projects, which can be executed faster and may deliver more reliable results.

DBA – Lean Practice Impact Evaluation – Industry Feedback

The first question addressed the potential to support the lean practice Quality Management by the DBA *Product Quality Monitoring.*

From the interviewee's perspective, data-based quality monitoring, such as visual inspection, has an enormous potential to automate quality inspection. Currently, a large part of visual inspection is done by a considerable number of employees. Due to the monotonous nature of the task, it is difficult to keep concentration high during the whole shift. Decreasing concentration increases the likelihood of defect products not being detected and sorted out. Camera systems, in contrast, promise a constant inspection quality. Also, visual inspection systems are faster and, although being extremely expensive to acquire, likely to be more cost-effective in the long run.

Nevertheless, until today, a camera system-based visual inspection also does not achieve a 100 percent accuracy rate. For products with complex geometry, the material handling effort is too high to justify automated inspection economically. Nevertheless, for identifying defect products, the impact of automated product quality monitoring is expected to be high. However, as sorting out defect products does not improve the quality of the product systematically, the general impact on the lean practice Quality Management is only moderate.

The second question addressed the impact of the DBA Predictive Maintenance on the lean practice Continuous Flow.

The argumentation that equipment stability supports flow, based on Womack and Jones (2003), was presented as justification for the initial assumption of a strong impact. According to the interview partner, the assumption that equipment stability supports flow can be backed from practical experience on the shop floor. Because Predictive Maintenance may reduce equipment downtime, the OEE and thus, the stability is positively influenced. Hence, Prescriptive Maintenance supports flow.

However, as technical breakdowns and maintenance are only two determinants of the OEE among others, the impact of Predictive Maintenance on Continuous Flow is more realistically to be rated as moderate than as high. Only if all other requirements for a high OEE, such as on-time material delivery are met, the impact of Predictive Maintenance on Continuous Flow is considered high.

The third question addressed the impact of the DBA Production Scheduling on the lean practice Pull/Kanban.

As discussed in chapter 4.2.3.3.2, literature documents a conflict of central IT-based production planning, promoting the push principle, and lean promoting the pull principle (Maguire, 2016, p. 32). The representative of Company B recognizes a theoretical conflict but argues that both approaches work well together in reality. In Company B's production sites, push and pull are used at the same time. The pull approach is used to trigger the production process of a product at the moment a customer places an order. The customer order is transferred digitally to the planning system, thus following the logic of an e-Kanban system.

The material flow between multiple process steps is then planned by the central planning system. Hence, Pull/Kanban is made redundant between the process steps. As a result, fewer WIP is necessary as Kanban inventory between process steps is not needed. Having a traditional LPS, with the Kanban system as the only mean of material flow management in mind, the example of Company B indicates at least a partial substitution of the lean practice Pull/Kanban by the DBA Production Scheduling.

5.2.3.8 Implications Case II: Company B

Table 24 summarizes the key implications of the second case study.

Table 24: Key implications – Case II

I	Company B - Automotive Supplier Company	P
I.1	Company B operates a company-specific LPS. A digitalization strategy exists with a strong focus on a cross-plant MES system. The digitalization strategy is not aligned with the lean strategy of the LPS.	134
I.2	Organizationally, the responsibility for lean and for digitization is separated. The lean unit is involved in projects to ensure the applicability and user-friendliness of IT solutions for the shop floor during implementation projects.	134 - 135
I.3	Company B already uses eight out of twelve DBAs at least partially in one of its sites. Two more DBAs are currently in the testing phase.	135
I.4	Company B has implemented a new MES system, comprising a Global Performance Cockpit to track performance metrics and generate early warnings of critical OEE levels.	136
I.5	Challenges for the MES system were the installation of sensors at all relevant equipment, definitions of interfaces to the MES system, and the standardization of metrics.	136
I.6	A key organizational challenge for introducing DBAs is change management. Employees need to understand the value of a new application to create acceptance for the DBA. Acceptance is threatened by extra effort for data collection as well as the fear of job loss.	136 - 137
I.7	Ongoing leadership commitment is critical for the implementation of long term projects.	137
I.8.	General technological challenges for DBAs are the integration of several data sources or IT systems, IT system stability, and data protection against data theft.	137
I.9	Data protection is increasingly important, to avoid misuse of personal related data as well as theft of manufacturing data by externals.	137
I.10	All employees working with a new system or application need to have basic training to understand its functions but also to understand the importance of error-free data that needs to be entered manually.	138
I.11	To drive DBAs projects, a holistic understanding of IT and manufacturing systems, as well as data analytics, is required on the management level. For the actual implementation, experts from both fields need to be involved in a cross-functional project team.	138
I.12	Full transparency of data and processes is considered as the main benefit of data utilization for lean.	138

I: Implications, P: Page

5.2.4 Case III: Company C

5.2.4.1 General Information

Company C is a major European ICT company and has around 20,000 employees generating a revenue of around $10 billion. It offers corporate and residential customers mobile and fixed-line telephony, internet, digital TV, and various cloud services. In addition, Company C is a large provider of IT services in the energy, advertising, entertainment, banking, and healthcare sector. The main contact during the joint project and the case study is the Head of Lean Management.

Company C constitutes a special case in this selection as it does not operate a manufacturing system in the sense of transforming raw material to finished products. Internally, however, the physical infrastructure to provide its services, including internet, telephony, and TV, is considered as the company's manufacturing system. As an ICT company, Company C has much experience in collecting, processing and analyzing data. Due to the high maturity in terms of data utilization, it is useful to include the company in the study even without a traditional manufacturing system.

As the DBAs introduced in chapter 4.1 are applicable within a manufacturing site only, they are not part of this case study. Instead, this case focuses on use cases of data utilization, which are different from those in the manufacturing area. Contrasting findings of this case study to the other three allows deriving conclusions whether particular challenges and enablers are specific to the manufacturing industry or can be generalized more broadly.

5.2.4.2 Lean Status Quo

"We are active in an increasingly competitive market where we need to address rising customer expectations, higher volumes, and changing technologies at the same time, all while lowering the cost base of our company—this means that we need to continuously improve our value creation activities."

With this statement, our main partner at Company C explains the criticality of CI and lean thinking in its organization. The company is guided by the five lean principles: define value from the customer perspective (1), identify the value stream (2), flow (3), pull (4), and strive for perfection (5). Customer orientation is one of three strategic objectives and customer satisfaction is a central metric for the company and relevant for financial bonuses for all employees.

Lean is understood as a system to create transparency on goals and value creation, to foster employee empowerment, and to support CI. The company emphasizes that lean should be carried out by all employees and managers in the organization. CI is requested and encouraged in different ways. As part of their agile work in development areas, participants meet every two weeks for retrospectives to critically evaluate the working method and the results of the past weeks to find ways to work smarter. Operating divisions, including network operation and customer service, use an app specifically designed to provide improvement suggestions. The app is embedded in a feedback process that ensures timely and qualified feedback.

Interestingly, from the lean practices of chapter 2.2.3, only the practices Quick Changeover Techniques and Lot Size Reduction are not part of Company C's lean efforts due to the lack of a physical manufacturing system.

Preventive Maintenance is used for base stations for mobile communication. For Quality Management, the quality of phone or internet connection is monitored. Continuous Flow aspires for customer tickets across different departments. The Pull/Kanban system is used to trigger the replenishment of warehouse products. Kanban is also used in software development as part of agile working methods. Value Stream Mapping is described as an “all-purpose weapon” of Company C. It is regularly applied in operating divisions as well as in software development, always aiming to identify non-value adding activities and reduce waste. Cross-functional teams are often set up as DevOps teams, primarily for product development. The practice Self-directed Work Teams is a key element of agile working often used in software development. The main objective of Self-directed Work Teams is an increased speed of decision-making.

The examples of the application of the lean practices show that Company C builds on traditional lean practices and has adapted the lean practices to a non-manufacturing environment and a modern working setup, including agile working methods.

5.2.4.3 Strategic Alignment of Lean and Digitalization

As of today, Company C has not formulated a dedicated digitalization strategy. However, Operational Excellence is one of three strategic objectives. Operational Excellence comprises the internal digital transformation, including an increased level of process automation (e.g., robotic process automation) as well as data analytics applications (e.g., ML applications). Operational Excellence aims to reduce the cost basis, to free cash flow for investments in new business areas.

Company C has a corporate lean unit, headed by the head of LM. Due to the fact that digitalization is happening parallel in many areas of the company, there is no single person responsible for digitalization. The internal digital transformation is driven by the process automation team.

Organizationally, the lean team and the process automation team are separated. There is no integrated leadership of lean and process automation, as both teams have very different tasks and core competencies. However, there is close cooperation between the two teams, which is mainly based on the fact that the value of each other is known. A regular exchange takes place, with discussions focusing on how existing cases can be handled most effectively: only automate them or design them lean beforehand and then automate them.

5.2.4.4 Data Utilization Use Case: Smart Support

Company C possesses a very high level of knowledge about disturbance patterns of its services. Thereby, the knowledge of the whole organization exceeds the experience a single service agent can possibly have. Exploiting this organizational knowledge to improve customer satisfaction of Company C's customer support was the main driver to develop the Smart Support system.

Customer satisfaction in this context is measured by two metrics. The first metric is the time a customer has to wait until he or she receives a solution to the problem. The second metric is the first-time right rate, which measures the ratio of the number of customer issues solved with a sustainable solution at the first contact with the service unit, divided by the total number of customer issues.

To support service employees in their daily work, the Smart Support system provides decision support. The system builds on an ML application that is trained by a database comprising all recorded quality issues in the past, along with successful mitigation strategies.

Smart Support works as follows. A customer reports per phone or internet a problem to the service employee. Based on the customer's problem description and a near-real-time measurement of the customer's connection, an ML application analyzes the problem and searches the database of past quality issues for similarities. If the problem description matches an existing problem and a respective solution is documented, the solution is proposed by the system to the service employee. If more than one solution may be plausible, the system ranks the suggestions according to the likelihood of a sustainable solution to the described problem.

Smart Support supports the service agent mainly in two ways. First, it enables the agent to propose a solution that is informed by all tracked quality issues that occurred and solutions proposed in the past. Therefore, the likelihood to find the best solution is increased strongly. Second, Smart Support reduces the number of interactions the service employee has to perform with its computer to identify a possible solution and hence, allows the employee to focus more on the customer.

In the aftermath of the communication with the customer, Company C tracks whether the problem is solved and whether it is solved sustainably. Both data points are feedbacked, thus enlarging the database and improving its accuracy every time the Smart Support system is used. As a result, blank spots of the system, where no data-based solution can be provided, are reduced over time.

Smart Support has considerably increased customer service by reducing the time until a solution is given and by reducing the number of non-sustainable solutions. By minimizing non-sustainable solutions waste is reduced and service employees use the freed time for other tasks.

Three technological challenges needed to be overcome to implement Smart Support. A key challenge was to translate the implicit experience-based knowledge of service employees to explicit machine-readable knowledge. As a result, in the beginning, only a small fraction of quality problems were covered in the database and the system needed time and permanent input of quality problems and solutions to increase this share. By now, the system covers almost all known quality problems and is applied in almost every interaction of a service employee with the customer. A second challenge was to move from the existing IT system to the new one without interrupting the customer support service. And third, a new process had to be designed for customer support. While the existing system required to follow a fixed path through the system, the new one allows jumping directly to a solution with the highest likelihood of solving the problem sustainably.

In the early phase of the Smart Support System, service employees rejected the idea of ML-based decision support. Some feared that the system would patronize them and expected a downgrade of their job from being a competent, self-determined solution developer to an order receiver that only communicates the system's solution to the customer. Company C took these concerns seriously and decided to design the Smart Support system in a way that always allows the service employee to take the final decision. If specialized knowledge enables him or her, the employee is free to propose a solution that differs from the one suggested by the system. The approach is comparable to a driver of an autonomous car who still has the power to oversteer the autopilot.

From a practical perspective, the system gets better every day as the probability of blank spots in the system decreases permanently. In the meantime, Smart Support has proven its value to enhance the job of service employees and increase customer satisfaction, thus resulting in a broad perception of the system as welcomed support. Nevertheless, keeping the service employee “in the driver seat” remains essential for the acceptance of the system.

5.2.4.5 Challenges and Enablers of Data Utilization

Technological Challenges and Enablers

Apart from the technological challenges described in the use case above, no further, more general technological challenges were discussed.

Organizational Challenges and Enablers

The main contact of Company C is Head of Lean Management, which explains his strong process orientation and ambition to enable flow within his organization. In an ICT company flow usually does not refer to physical but virtual entities; for example, customer tickets. To enable the flow of virtual flow objects through the organization, a standardized definition and identification of theses flow objects is required as well as access to all relevant data for all involved departments. Ensuring a standardized definition and a unique identification of virtual flow objects within the whole organization is considered a major organizational challenge.

Company C has defined the objective of becoming a data-driven company, hence, to learn from data to support operational and strategic decision-making. However, Company C has recognized that transparency created by data is not universally welcomed. Transparency might also be perceived as an instrument to control individual performance of employees. This perception results in a negative attitude towards data collection and analytics among employees.

The company does not deny that data is increasing transparency, also including the possibility for objective performance evaluation. The message, however, is that data is not used to control individual behavior but enables a transparent and fair leadership based on objective metrics. A transparent communication, which metrics are collected and used for performance assessment, has helped to increase the acceptance for metric-based performance assessment among employees.

Employee Challenges and Enablers

Employee qualification is an essential enabler for applying new DBAs such as the Smart Support system. The optimal qualification thereby depends on the role of the employee. End-users benefit from training that not only focuses on how to use a new application but also conveys an understanding of the underlying logic of the application. Understanding the underlying logic has proven to be helpful to use the application more effectively, compared to only follow the user guidelines. However, the effort for additional training has to be economically reasonable. Advanced technological knowledge (e.g., of the ML algorithm) does not provide additional value for the daily job of service agents and thus is not part of the company provided training.

To develop and implement applications such as the Smart Support system, developers and managers require advanced skills in various areas, such as ML, data science, and DM. To cover the need for qualified employees, Company C has created an advanced training system building on two pillars.

The first pillar consists of formal education in cooperation with a well-known university. In a 150-hour course, comprising web-based and classroom training, more than 100 employees are trained as a data scientist. Although Company C offers a three-digit number of places in the course, the demand is exceeding the number of available places by far, demonstrating the high motivation of employees for advanced training in the field of data analytics.

The second pillar consists of a systematic building and the exchange of internal know-how. One concept of the second pillar is called Stage. During a predefined period, often between 20 to 40 weeks, an employee from a non-IT unit works for a fixed percentage of his time in an IT department or data science team. With this form of learning on the job, cross-functional know-how exchange is fostered as well as the creation of personal links between different departments. Nevertheless, finding and training sufficient data-affine employees internally is very challenging as the interview partner admits: “We know that we need more data-affine employees.”

To complement internal know-how building, Company C regularly screens the labor market for qualified persons. However, well-trained people with an affinity to data are rare and many companies compete for them on the market. Although a strong company brand is helpful to convince data talents they know their market value, resulting in comparably high salaries. Therefore, Company C is active to hire data talents, however not at all costs.

5.2.4.6 Impact of Data Utilization on Lean

“Gaining direct data-based insights on where the biggest potentials for optimization and the next improvements are, as opposed to relying on opinions and prioritizing them” was the answer of a second member of Company C's corporate lean team to the question of what the next step in lean regarding digitalization will be. The quote summarizes the intention of Company C to use data to identify improvement opportunities and support decision-making.

Data-based support of optimization is a pivotal element of Company C's vision for the next stage of lean. The why for optimization should thereby be determined by business needs. Data and advanced analytics then provide insights on what and how to best improve.

Also, the company strives to use DBAs, such as the Smart Support System, to increase customer value while reducing waste and thus following two lean principles. Robotic process automation is used in combination with ML to automate processes. For instance, an intelligent chatbot was developed to minimize response time to customers for written requests (best case less than one minute). Voice recognition is another ML application that is tested to reduce the effort for customers to identify themselves on the telephone.

Most value-creating processes of Company C are non-physical. Data collection and analysis allows the monitoring of the performance of these virtual processes. Comparable to traditional manufacturing KPIs, virtual process performance metrics increase transparency on process performance and indicate occurring, but invisible, process problems.

In conclusion, as Company C does not have a physical manufacturing system, it is not feasible to evaluate the impact of data utilization on traditional lean practices. However, on the level of lean principles, the potential of data utilization to support lean objectives, especially CI, customer value creation, and waste reduction is evident.

5.2.4.7 Implications Case III: Company C

Table 25 summarizes the key implications of the third case study.

Table 25: Key implications – Case III

I	Company C - Information and Communication Technology Company	P
I.1	Although being an ICT company without a physical production system, Company C follows the five lean principles and applies several lean practices in their daily operations.	142
I.2	Lean and digitalization are not managed in an integrated way. However, the skills and competencies of the lean team are known to the teams driving the digital transformation and vice versa. This knowledge fosters a close collaboration on a project basis.	142
I.3	Company C is very advanced in data utilization. The Smart Support system is a mature ML application that uses historical data to increase customer satisfaction.	143
I.4	A key technological challenge of the Smart Support system was the translation of implicit knowledge of service employees to explicit knowledge eligible for ML.	144
I.5	An initial barrier was the perception of the new system as a threat to the autonomy of employees and their self-image as a solution developer. By granting final decision-making power, also against the proposition of the system, the perception of the system has changed and it is now considered a useful supporting tool in customer services.	144
I.6	Standardizing definitions and identifications of virtual flow objects, such as customer tickets, across all departments is a major organizational challenge.	145
I.7	Increased transparency due to data collection was partially perceived as an instrument for controlling employees. Transparent communication about which metrics are used for performance evaluation and which are not has helped to overcome this perception.	145
I.8.	Employees require job-specific training to use DBAs effectively. End-users of an application, such as the Smart Support application, need to be able to operate the system and understand the underlying logic behind the application.	145
I.9	Developers and managers need advanced training, which is realized with a dedicated external course on data analytics as well as with learning on the job in different departments. Demand for advanced training is currently exceeding the offer by far.	146
I.10	Data talents are attractive for many companies, thus making it difficult and expensive to hire data scientists from the labor market.	146
I.11	"Gaining direct data-based insights on where the biggest potentials for optimization and the next improvements are" is the primary value proposition of digitalization for the next step of lean.	147
I.12	Company C uses data in several ways to support the lean objectives of customer value and waste reduction, including several advanced applications of ML.	147

I: Implications, P: Page

5.3 Academic Expert Interview

The following two chapters present the results from expert interviews with two senior academics. Expert interviews were conducted for two reasons. First, to broaden the basis of qualitative data for the cross-case analysis in chapter 5.4. As described above, the experts serve as "surrogates for a wider circle of players" by contributing insights from several DBA projects in the manufacturing industry. Second, the senior academics were asked to challenge a theoretical model that emerged during this dissertation project. Due to their practical and theoretical experience in the field of data utilization in manufacturing, they are deemed to be competent sparring partners to discuss the initial versions of a DBA value model. The DBA value model and the derived ROI dilemma of DBAs theory will be introduced in the consolidation chapter 6. Feedback given during the two expert interviews on the model and the theory is not discussed in this chapter but instead in the respective chapter 6.3.1

Both academics have an excellent reputation in the field of manufacturing data analytics and share the perspective that doing research in manufacturing data analytics requires close interaction with real-world manufacturing companies. The guideline used to structure the semi-structured interview is based on the case study interview guideline but has been adapted to the particular research focus of the interviewee (see Appendix D).

The two interviews share a brief introduction, providing background information about the position, career, and research focus of the interview partner. The second section investigates the motivation of companies to initiate DBA projects. The third section is distinct. Expert I is currently involved in several data utilization projects and is, therefore, asked to provide some insights on the current status quo of selected DBAs. Expert II has lived, worked, and conducted research for several years in the USA and Europe. Due to the fact that he knows both regions well, he was asked to evaluate differences between the USA and Europe regarding the application of ML/AI.[27] The fourth section focuses on the main challenges and enablers for data utilization in manufacturing. The last section summarizes key implications per interview.

5.3.1 Expert Interview I

5.3.1.1 Background Information

Dr. Guido Schuster is a professor at the University of Applied Sciences Rapperswil in Switzerland and works at the Institute for Communication Systems (ICOM). His professional career combines research and industry experience. He has gained a master's and a doctoral degree at Northwestern University, Illinois, USA. During his

[27] The term machine learning is used synonymiously with the term artificial intelligence

Ph.D., he worked for the Motorola Corporate Research Laboratories in Illinois and was part of the development of the MPEG-4[28] standard.

In the following years, Professor Schuster was involved in several research projects. He specializes in applied science and development in the following two areas: First, digital signal processing, including computer vision and image processing; and second, AI, including ML and deep learning. He has been honored with several awards for his research and holds more than 60 international patents. As a researcher, he has published more than 65 peer-reviewed publications.

5.3.1.2 Motivation of Data Utilization in Manufacturing

Based on his experience, Professor Schuster argues that initiatives for data utilization in manufacturing are often initiated by the top management, although he has observed a lack of technical expertise among many top executives. One reason for this is that management is afraid that their company is lacking behind its competitors in terms of data utilization capabilities and therefore pressures the organization to search for opportunities for data utilization. Today, stakeholders of larger companies, such as the supervisory board, expect the top management to have a kind of “data strategy.” Therefore, data-based initiatives are regularly driven top-down.

Comparably few initiatives are driven by manufacturing responsible. It appears that they are, in general, already quite satisfied with the current situation, and as a consequence, have little motivation to invest much effort to collect data or to lay the technological foundation of DBAs. Except for large and technological mature companies, the current availability and quality of data of manufacturing companies are quite limited. In addition, Professor Schuster perceives some resistance by production managers and shop floor employees against DBAs. Shop floor employees may perceive DBA as a driver of rationalization and finally as a driver of job reduction, resulting in little enthusiasm to support the introduction of such applications. Production managers are often not eager to initiate DBAs for a different reason. In the experience of Professor Schuster, they tend to feel criticized if somebody suggests they use data to support decision-making instead

of relying on the personal experience of the production managers.

In conclusion, in many cases, top management initiates projects to implement DBAs to demonstrate their willingness to prepare the company for the digital transformation, whereas manufacturing units tend to show initial inertia. Overcoming the initial inertia is a crucial enabler for the successful introduction of DBAs.

[28] MPEG-4 is a method of defining compression of audio and visual (AV) digital data.

5.3.1.3 Status Quo of selected DBAs in Manufacturing

Due to his broad experience in data utilization in manufacturing, both from an academic perspective but also from several industry projects, Professor Schuster was asked to comment on the current status of DBA projects he was or is personally involved in.

DBA Production Scheduling

The objective of the DBA Production Scheduling is to determine the optimal production plan, based on current orders and the availability of manufacturing resources. Professor Schuster is currently involved in a production scheduling research project with a large automotive supplier.

The project seeks to optimize production scheduling to reduce the number of changeovers of two machines. Previously, production scheduling for the two machines was done by intuition without the support of operations research optimization tools. By describing the current situation as a mathematical optimization problem and solving the problem with a brute force approach, the productive time of both machines was increased by more than three percent, thus resulting in higher daily throughput.

Professor Schuster reports a gap between operation research in the academic world and the industry. While academia deals with highly complex operation research problems, most industrial companies fail to formulate and solve even comparably simple operation research problems. In addition, he has identified a lack of a strategic approach to use optimization applications systematically. Instead, most companies follow their instinct for production scheduling. As only a few percentages of improvements (e.g., of equipment utilization) already yield considerable savings, systematic, data-based production planning promises high potentials for cost reduction, or as Professor Schuster puts it: “There is still money on the street.”

DBAs Condition Monitoring and Predictive Maintenance

The objective of a current project on Predictive Maintenance using ML is to monitor the current condition of the equipment (corresponds to the DBA Condition Monitoring) and to predict the remaining useful life (corresponds to the DBA Predictive Maintenance). The value proposition of sensor-based condition monitoring is to be able to monitor several signals permanently and simultaneously. The intention is to use internal signals to identify failures of the production equipment before they negatively impact product quality.

The task of Professor Schuster in this project is the analysis of selected sets of data to identify patterns within the data. The project has made good progress in achieving the first objective, which is to monitor the current condition of the machine. However, a substantial barrier emerged that is currently hindering the project from making progress in regards to predicting the remaining useful life. Because the company pays much attention to maintain the machine properly, only very few failures occur. However, sufficient data on past failures are essential to finding patterns within the data. Without failure data, prediction of future data proves to be extremely challenging.

A second yet related problem is that the lack of failure data hinders the research team to validate the predictions made. Currently, the machine is running with very few disturbances, but nobody can seriously tell whether this is an effect of predictive maintenance or if it would be the same with less maintenance effort. Although not backed by data, Professor Schuster holds the view that currently, more effort in maintenance is invested as economically reasonable. Furthermore, based on this experience, he is generally rather pessimistic about whether predictive maintenance can fulfill the high expectations raised by scholars and practitioners.

DBA Material Flow Management

A third project is concerned with optimizing the flow of materials within an intralogistics system. The objective is to increase the throughput (number of completed orders/time), by reducing the number of sorting operations. The optimization strategy includes the use of simulation software. Therefore, the situation was formulated as a material routing problem. Iterative simulation is selected over a long time planning for two reasons. First, planning a longer time horizon makes the problem increasingly complex and therefore requires much time and computing power. Second, problems (e.g., machine breakdowns) may occur dynamically, thus making long-time plans obsolete. To find the best solutions, different material routing scenarios are implemented and simulated. By comparing the simulation results, the best performing scenario is determined.

Due to the limited complexity of the simulation problem and the simulation model, the technological requirements have not been a barrier. A modern standard personal computer was able to solve the simulation problem within hours. However, as the complexity of the problem is increasing exponentially when adding new elements, Professor Schuster expects that in the near future, simulations will be run on cloud computing services using the scalable computing power. More complex simulation problems also require high-performing network connections.

However, similarly to the status quo of the DBA Production Scheduling, the manufacturing industry seems to be far from exploiting the full potential of simulations to optimize material flow. The application of simulation techniques is still in its infancies in most manufacturing companies. As a result, the discussion about the *optimal* material flow is more an academic than a practitioner issue.

5.3.1.4 Challenges and Enabler for Data Utilization in Manufacturing

Technological Challenges and Enablers

Key technological challenges identified by Professor Schuster are well in line with the key requirements identified from the literature and summarized in chapter 4.1.6. A basic challenge for DBAs is to provide the required data. Manufacturing equipment needs to be equipped with sensors and connected to a central data management system. Depending on the application and the complexity of the problem, the bandwidth of data connections, connection stability and access to sufficient computing power are additional requirements.

Organizational Challenges and Enablers

According to Professor Schuster, a major challenge for companies to introduce DBAs is a lack of technological expertise, especially on the higher management levels. As a result, the decision if a management team appreciates or rejects DBAs is often "a question of faith" rather than the result of an informed decision. From personal experience, Professor Schuster has observed that managers tend to follow his opinion without being able to challenge it critically.

However, a certain level of technical expertise of company representatives is essential to serve as a critical sparring partner for consultants and service providers. Having this minimum level of expertise within the company is one of the traits that distinguishes companies that are successful in implementing DBAs from companies that tend to fail to do so.

Based on long-term experience from several industry projects, Professor Schuster has identified the following two enabling factors for implementing DBA projects.

First, DBA projects need to be driven by a dedicated person, that has the time and the competence to manage and push the project. DBA projects are too time-consuming to be driven in addition to the day-to-day business. Also, the person optimally has access to financial and personal resources within the company. This is linked to the requirement of having a certain degree of seniority, on one hand, to have access to the top management and on the other hand, to overcome internal resistance. Finally, a certain "terrier" mentality and persuasiveness are also useful to overcome internal resistance and to keep stakeholders motivated, in case the project is delayed or does not deliver as intended.

Second, top management support is critical. A key barrier to implementing DBAs is that nobody can predict the result or the value of the applications in advance. Therefore, management commitment is inevitable, especially if DBA projects do not deliver the expected results in the first attempt. Management buy-in is therefore critical for the implementation, as quick wins are rarely found. Successful DBA companies show persistence against setbacks in the early phase of DBA projects.

The combination of high technical expertise and high executive power is found in organizations that have a CTO. These organizations are found to have an advantage for implementing technologically challenging projects, to which some DBA projects surely belong.

Professor Schuster regularly works as a consultant and solution provider with companies. From his perspective, members of the middle management with technical expertise have so far been the best project partner as they tend to have some seniority in the organization but also the time to be involved personally in the project.

When working as a consultant for DBAs, another very fundamental challenge becomes apparent, which can briefly be characterized as a “chicken and egg problem.” Company managers typically like to have an estimate of the return on investment (ROI) before granting resources for projects. However, an ROI estimation of DBAs that include data analytics called data analytics DBA (see chapter 4.1.5.3) is highly afflicted with uncertainty. Finding useful patterns in the data cannot be guaranteed seriously before the analysis has been performed.

Without having a clear business case, however, managers are very hesitant to invest in data collection. At the same time, the consultant can only demonstrate the value of a data analytics DBA, after data was collected and analyzed.

The problem can be mitigated if the consultant can provide successful practice examples from other, but similar, situations. Having seen a proof of concept often increases the willingness of managers to take the risk to invest resources in a project without a predictable payoff.

Employee Challenges and Enablers

Having discussed the high relevance of technical expertise regarding data utilization to drive DBA projects, questions arise regarding how this expertise can be built most effectively. A full content course on data analytics would be highly time-consuming and would require a lot of existing knowledge in the fields of mathematics and programming.

Assumingly, such a course would not be selected by many top managers, as they neither have the time nor the basic knowledge required. However, a tailored course covering the basics, to enable managers to evaluate internal and external DBA project pitches adequately, would be valuable for many companies. Nevertheless, at the time of the interview, the University of Applied Sciences Rapperswil had no plans to design and offer such a course to top executives.

5.3.1.5 Implications Expert Interview I

Table 26 summarizes the key implications of the first expert interview.

Table 26: Key implications – Expert interview I

I	Expert Interview I	P
I.1	Data utilization projects are often driven by top management rather than by manufacturing managers. Companies' stakeholders expect the company to have a strategy to utilize data.	150
I.2	Manufacturing units tend to show some internal resistance. Some manufacturing managers feel criticized by the request to use data analytics as a basis for decision-making. Shop floor employees may perceive DBAs as a threat to their jobs.	150
I.3	Overcoming the initial resistance of manufacturing employees is a crucial enabler for the successful introduction of DBAs.	150
I.4	There is great potential for cost savings for manufacturing companies by switching to a systematic, data-based production scheduling, using operations research methods.	151
I.5	A key barrier to predictive maintenance applications may be the lack of failure data to detect patterns or to validate predictions. Well-maintained machines produce too few errors for analyzing the data for patterns.	151
I.6	High complex simulation problems require scalable computing and high performing data connections. However, the application of simulation techniques is still in its infancies in most manufacturing companies.	152
I.7	Technological challenges identified in praxis are well in line with those identified in the literature.	153
I.8.	If DBA projects are driven by a dedicated person, with access to sufficient financial and human resources and the ability to overcome internal resistance, the success rate is significantly higher. Middle managers, with technical expertise, are suitable partners for DBA projects.	153
I.9	Top management support is critical for DBA projects, especially in case of setbacks in the early phase of a DBA project.	153
I.10	The chicken and egg problem applies to data analytics DBA projects. The management asks for guarantees of results before granting resources for data collection and analysis. However, only after the data analysis, the value of the project can be determined.	154
I.11	A tailored course for managers, covering basic concepts of data analytics, would be useful to enable them to evaluate proposals for DBA projects adequately.	154

I: Implications, P: Page

5.3.2 Expert Interview II

5.3.2.1 Background

Dr. Thorsten Wuest serves as Assistant Professor for Smart and Advanced Manufacturing at the West Virginia University and works at the Department for Industrial and Management Systems Engineering (IMSE). Among his research focus is SM, ML, data analytics, hybrid analytics, as well as information and data management with a focus on manufacturing systems. His research is characterized by an emphasis on an interdisciplinary and holistic approach towards analysis and optimization. He has published over 100 peer-reviewed articles. He is a member of the Editorial Board of the Journal of Manufacturing Systems (JMSY).

5.3.2.2 Motivation of Data Utilization in Manufacturing

Usually, projects to exploit data in manufacturing are initiated by top management. "We need to do something in regards to Smart Manufacturing" is what many leadership teams believe. However, the project ideas are usually rather unspecific, and it is the responsibility of the manufacturing unit to define, specify, and implement appropriate data utilization opportunities. Professor Wuest has seen a variety of reactions of the manufacturing unit to data utilization plans of top management. While some appreciate the initiative and soon become an active driver of the initiatives, some manufacturing units are quite resistant.

Generally speaking, data utilization is promoted from two perspectives. On one hand, manufacturing companies seeking to use data utilization to improve the performance of their production system. On the other hand, production machinery manufacturers develop and offer increasingly sophisticated applications building on data utilization. They perceive data utilization as an enabler for new business models, including pay-per-use contracting or outcome-based contracting, as well as an additional feature of their products, which serves as a sales argument.

From a financial and technological perspective, a key driver for the increase of data utilization in recent years is the development of reliable, yet inexpensive technical components. The performance and robustness of components, such as sensors, has significantly increased. At the same time, the cost of sensors has been reduced by 50 percent and even higher cost reduction applies to other components, including computing power, storage space, and bandwidth. Many companies have introduced cloud solutions that allow fast, reliable, and secure access to data. Also, algorithms and tools for data analytics have been refined and become more user-friendly. In conclusion, the ratio of required financial investments and expected outcomes has become attractive for a significant number of companies in recent years for the first time.

5.3.2.3 Machine Learning in the USA and Europe

According to the media,[29] the USA and China are the ML superpowers. Professor Wuest is an expert in the field of ML and has lived and worked for years in Europe and the USA. For that reason, he was asked to evaluate the differences between Europe and the USA regarding ML.

The positioning of the USA and China at the forefront of ML is correct. From a general perspective, the two nations are leaders in research and application of ML by far. The reasons for this development are multifaceted. One important enabler is the availability of an abundance of data. In the USA, large tech companies such as Google and Amazon collect enormous amounts of social media and customer data. Other companies such as UBER collect vast amounts of position data. The companies not only collect these data but invest large amounts of time and financial resources in exploiting it. In Europe, in contrast, comparably few companies collect data on such a large scale. Furthermore, data protection laws in Europe are far more strict than in the USA, thus reducing the opportunity for data collection and analytics.

Another reason is the ability of the USA to attract many of the leading researchers in the field of ML. Besides the large tech companies, governmental institutions such as NASA and military research institutions have the financial power to provide attractive working conditions for the best ML specialists. However, ML is not only driven by large companies and governmental organizations but also by agile start-ups. The USA has managed to established an active start-up community, combining technical expertise, entrepreneurial spirit, and high investments of venture capitalist firms. So in general, Europe indeed is already lagging behind the USA in ML and there is little evidence to assume that this will change in the future.

However, when focusing on the manufacturing industry, the picture is different. The applications of ML in manufacturing are often very specific and require a deep understanding of the manufacturing processes. Companies from Switzerland and Germany are often global technological leaders in their industries and therefore have excellent technical expertise. From personal experience, Professor Wuest argues that many European manufacturing companies are leaders in ML for specific applications. As a result, the large gap between the USA and Europe that is very evident for ML in general, is currently not recognizable in the manufacturing industry. Thus, as of today, ML is not a threat to the competitiveness of European manufacturing companies. Nevertheless, prospectively U.S. companies but also Chinese companies may increasingly benefit from the high maturity of their nations in ML.

[29] Frankfurter Allgemeine Zeitung GmbH (2017), ZEIT ONLINE GmbH (2018), Handelsblatt GmbH (2019)

5.3.2.4 Challenges and Enablers for Data Utilization in Manufacturing

Technological Challenges and Enablers

Building the technological infrastructure, including sensors for data collection, stable network connections for data transfer, cloud or on-premise solutions providing scalable capacity for data storage and computing power for data processing, has been a central challenge until recently. However, due to the sharp reduction in prices, the technological infrastructure is available for the first time for a large number of companies at a reasonable price. Therefore, the role of the technological infrastructure as a critical challenge will decrease.

As an expert for ML, Professor Wuest reports a challenge specific to ML in manufacturing that has not been discussed hitherto. Currently, a majority of available ML algorithms are not specifically designed for an application within manufacturing but instead are optimized for very large samples and data with rather low complexity, such as the recognition of images. In manufacturing, however, production data are often complex and have smaller sample sizes. While a social media ML applications can use billions of pictures from the internet for training, data to train an ML algorithm for predictive maintenance is limited to the data the machine generates itself.

Organizational Challenges and Enablers

Although companies have built some experience with DBAs, the fact that the result of DBAs on the bottom line is still challenging to estimate remains a key challenge. Answering the question "what is the real value generated by data-based applications" is still almost impossible to answer before the application was implemented and tested. There are many successful DBA examples, but also many examples without any sustainable impact. Even more important to inform investment decisions but also even more challenging to answer is the question "if we invest now, when do we benefit from the investment and how can the return on investment be measured?"

Although the success of a data-utilization project cannot be guaranteed beforehand, Professor Wuest was asked to describe patterns of companies that have a higher chance to implement DBA applications more successfully than other companies.

First, manufacturing companies intending to implement ML applications require a certain minimal level of digital maturity, including a certain level of digitalized processes and IT knowledge. For analog companies, the way to benefit from ML applications may be too tedious. Without the perspective of success stories, motivation, and management support, it is challenging to go full force all at once.

Second, the personality of the project owner is very important. Especially with more complex data analytics, DBAs require persistence to overcome barriers along the way. The project owner must have access to sufficient resources to drive the project. The same applies to project employees. Without the necessary technological resources (e.g., access to the data in the cloud) and time resources, the motivation of the employee and thus the likelihood of a successful DBA project decreases considerably. Providing enough resources is a management responsibility.

DBA projects often take more time than initially planned, or fail to deliver the promised results in the first iteration. To maintain the support of the management and to keep the motivation of all DBA project stakeholders high, Professor Wuest recommends built-in Quick Wins from the beginning. Quick Wins are easy to achieve intermediate results, which are not necessarily part of the final solution but help to maintain the motivation of stakeholders on the way. An example of a Quick Win would be a threshold-based control system as part of a predictive maintenance DBA.

Third, maybe biased to his role as senior academic, Professor Wuest recommends companies to partner up with partners to use external resources. Universities may serve as an attractive collaboration partner for manufacturing companies for three reasons. First, Master and Ph.D. students have the capacity to work dedicated on a project, while company employees usually have other obligations in parallel. Second, Master and Ph.D. students are intrinsically motivated to complete a project within a given period. And third, universities have access to valuable skills and resources while being less expensive than consultant companies.

No clear answer could be given to the question of the main reasons for the failure of DBA projects. However, a phenomenon that was frequently observed by Professor Wuest was that companies started to collect data without having defined explicitly for what they intend to use the data in advance. The lack of a clear predefined use case then resulted in data collection and data analytics, both consuming resources, but no practical application. Even though data collection is getting cheaper and more convenient, Professor Wuest recommends starting with understanding the problem, derive a use case for data utilization, and only then start to collect and analyze data.

Employee Challenges and Enablers

As SM technologies and data analytics are likely to change the way of working in the manufacturing industry, the required skill sets of employees will change accordingly. In the ideal situation, manufacturing employees combine manufacturing know-how and data analytics expertise. Professor Wuest assumes that the combination of both skills will be rather a result of the training of technicians and manufacturing engineers in data analytics skills than the training of data scientists in manufacturing know-how. As technicians and engineers are usually technology and mathematics affine, they have good preconditions to learn data analytics skills quickly. The starting point for successful training is thereby an intrinsic motivation for additional qualification. Companies are advised to conduct hands-on workshops with actual use cases to demonstrate the opportunities of data utilization. Professor Wuest recommends a step-by-step qualification roadmap, where the employee decides self-reliably which level of qualification is at the moment most appropriate for the current tasks. Companies should focus their effort on fostering employees' motivation for additional qualifications; for example, by providing the perspective of more interesting jobs.

5.3.2.5 Implications Expert Interview II

Table 27 summarizes the key implications of the second expert interview.

Table 27: Key implications – Expert interview II

I	Expert Interview II	P
I.1	Data utilization projects are mostly initiated by the top management, which feels pressured to “do something in regards to Smart Manufacturing.”	156
I.2	The manufacturing unit is responsible for identifying and implementing suitable applications. While some manufacturing units are actively supporting data utilization, others show strong resistance.	156
I.3	The emergence of data utilization in manufacturing is driven by production machinery manufacturers looking for new business models, and production machinery users seeking to improve the performance of their production system.	156
I.4	Components of the technological infrastructure needed for data utilization have become significantly better and cheaper in recent years, thus making it economically attractive for companies to invest in data utilization for the first time.	156
I.5	The USA and China are technology leaders in AI and ML, with Europe lagging behind. However, in the manufacturing realm with specific requirements, European companies are competitive in terms of AI utilization	157
I.6	A technological challenge specific to manufacturing companies is to adapt the existing ML algorithms from other areas of application to the specific conditions of manufacturing.	158
I.7	A key challenge for data utilization projects is to estimate the value prior to the project. Forecasting the ROI and the time of the ROI seriously is currently almost impossible.	158
I.8.	A certain digital maturity of a company is a prerequisite to implementing data utilization applications successfully.	158
I.9	The owner of a data utilization project needs to be persistent and equipped with sufficient resources.	158
I.10	Data utilization projects often take more time than initially planned. Built-in “Quick Wins” is an option to keep the motivation of project stakeholders and management high.	159
I.11	New technologies require additional qualifications. Companies should provide opportunities for training to achieve a self-determined *optimal* level of qualification.	159

I: Implications, P: Page

5.4 Cross-case Study Analysis

After having presented three case studies as well as two expert interviews with academics, this chapter seeks to draw generalizable conclusions based on a cross-case analysis (R. Yin, 2009, p. 49).

According to Eisenhardt (1989, p. 541), the idea of cross-case analysis is to "force investigators to go beyond initial impressions, especially through the use of structured and diverse lenses on the data." The tactic is motivated by the fact that people and thus also researchers are poor in information processing. Eisenhardt (1989, p. 540) argues that humans tend to jump to conclusions prematurely based on a limited set of data, disregard statistical properties, overlook conflicting evidence, and are over proportionally influenced by elite respondents. These information-processing biases put researchers at risk to draw premature and even false conclusions. Looking at the data in divergent ways and contrasting evidence from different cases is a good strategy to counteract these biases. Voss et al. (2002, p. 215) agrees and emphasizes that seeking confirmation from using multiple data sources is important in case study research as it leads to more reliable and more generalizable results.

Eisenhardt (1989, p. 540) and Voss et al. (2002, p. 214) propose to define categories and then search the case material for similarities and differences between the cases. Categories are typically suggested by the research problem or by existing literature. Contrasting seemingly similar cases, searching for differences may allow researchers to overcome too simplistic frames while searching for similarities in apparently different cases increase the problem understanding (Eisenhardt, 1989, p. 541). The results of the cross-case analysis are used for both testing existent theories and hypotheses as well as for developing or extending theory (Voss et al., 2002, p. 216).

This chapter contrasts and combines the findings from the case studies and the expert interviews and comprises three subchapters. Chapter 5.4.1 focuses on the current status quo of lean management in the case companies and investigates the question of whether these companies have started to integrate lean and digitalization from a strategy perspective, an organizational embedding perspective, or both.

Chapter 5.4.2 then focuses on data utilization in the industry. It describes the principal motives of companies to invest in data utilization and outlines the current status of the actual use of DBAs. A focus is set on collecting, contrasting, and consolidating challenges and enablers of data utilization from the cases and interviews. Chapter 5.4.3 contrasts the findings of the three cases to identify commonalities and differences between the manufacturing companies and the ICT company.

The cross-case study analysis builds on the implications drawn from the three case studies and the two expert interviews. To ensure transparency on which basis the conclusions are derived, the respective implications from chapter 5.2 and chapter 5.3

are referenced. For instance, the reference (C1-I.1) refers to the first implication of case I and the reference (E2.I.3) refers to the third implication of the second expert interview. For the readers' convenience, the references contain a hyperlink to navigate to the respective implication chapter. Furthermore, the implication overview tables contain references to the respective section where the implication is derived from. If findings from the qualitative studies are used in this chapter but have not listed as a key implication, direct references are provided to the respective section (e.g., p. 122).

5.4.1 Lean and Digitalization

The first section of the cross-case analysis is concerned with the current status of LM and the integrated consideration of lean and digitalization.

Assessing the status quo of lean is relevant for this dissertation for two reasons. First, as lean is understood differently by many researchers and practitioners, it is necessary to ensure a shared understanding of lean (e.g., is lean considered the foundation of the production system or even as a guiding principle of the whole organization, or is it instead perceived as a toolbox for the shop floor?). Second, only if the case study partners are familiar with the lean practices, and their company applies those practices, are they qualified to contribute to the DBA – lean practice impact assessment. The status quo of lean in the case companies is summarized in chapter 5.4.1.1.

Second, several researchers have found indicators that LM and digitalization is positively associated. Küpper et al. (2017) report higher potentials for cost savings of an integrated lean and industry 4.0 approach, compared to standalone approaches. Tortorella and Fettermann (2018) found in a quantitative study a positive relationship between LM and industry 4.0 technologies.

These findings are supported by two research contributions of the author of this dissertation. First, the survey presented in chapter 3 shows that a majority of 95 percent of the participants expects a mutually beneficial relationship between lean and digitalization (Macuvele et al., 2018). Second, Lorenz et al. (2019) found that companies with a high digital maturity and a high lean maturity tend to perform better than companies with only a high digital maturity or a high lean maturity.

These findings raise the question of whether companies share the perspective that combing lean and digitalization is beneficial and whether companies try to foster the integration of lean and digitalization. Therefore, chapter 5.4.1.2 summarizes the case study results on the strategic or organizational integration of lean and digitalization.

5.4.1.1 Status Quo of Lean Manufacturing

Company A and Company B operate a large manufacturing network. Both companies have organized their production according to lean principles and describe their system as an LPS. However, the production systems are company-specific and no 1:1 replica of the TPS.

Company A's LPS consists of nine elements and seeks to support the three competitive priorities of low cost, high quality, and high delivery reliability. More high-level objectives are customer and employee satisfaction. Company B has designed an LPS with a strong focus on CI, having the PDCA cycle at its heart. Employees are strongly encouraged to participate in CI by submitting improvement suggestions. Company C does not operate a physical production system and therefore has no LPS. Nevertheless, Company C works according to the five lean principles. Customer orientation is a key metric and a share of employees' salary is linked to achieving customer satisfaction-specific goals. Company C demonstrates that principles originating from LM can be transferred and applied successfully in non-manufacturing settings.

From an organizational perspective, all three companies have a central lean team that supports the whole organization with tools and methods but also training. Company A and B additionally have dedicated lean managers on a site level.

The 10 lean practices, identified by Shah and Ward (2003) and complemented by the author of this dissertation, are indeed widely established lean practices. Both manufacturing companies, Company A and Company B, apply all 10 of the practices at least in some of their plants. Although the degree of implementation varies between the practices, all 10 have proven to be relevant to modern manufacturing companies and are well-known to production managers.

Interestingly, even Company C has started to apply eight of the 10 practices, despite being an ICT company. Some practices, as preventive maintenance, are applied in the traditional understanding of the lean practice. Other practices have been adapted to non-manufacturing areas. For instance, the lean practice Continuous Flow is implemented to reduce the lead time of customer tickets, which are virtual objects that need to be processed by different departments.

In summary, for the sample of the three case companies, lean is of fundamental relevance as the underlying and guiding principle of the production system but also as a toolbox of lean practices.

5.4.1.2 Integrated Consideration of Lean and Digitization

As described in the introduction to this chapter, several publications suggest that the integrated consideration of lean and digitalization increases the effectiveness of both approaches mutually.

The assumption that the integrated consideration of lean and digitalization is beneficial is supported by the findings of the case studies. Company A explicitly characterizes its company-specific LPS as the basis for innovation (C1-I.1). Lean thinking is essential to define clear standards and interfaces. In turn, these standards are essential prerequisites for connecting hitherto distinct manufacturing units and thus pave the road to connected manufacturing and connected logistics. Company B highlights the importance of bringing in a lean perspective, as the ideal process from an IT perspective does not necessarily equal the perfect process from an employee perspective on the shop floor. To balance both requirements, lean and IT responsible therefore need to interact closely (C2-I.2). Quite similarly, Company C promotes the interaction of the lean team and the team responsible for the internal digital transformation. Before a business process is automated, both teams discuss if the current process is already suitable for automation, or needs to be designed according to lean principles first (see page 142).

In summary, all three companies argue concordantly for the integrated consideration of lean and digitalization activities. This finding motivated two questions. The first one is whether the desired close interaction of lean and digitalization is fostered by an integrated strategy that considers both approaches as complementary elements. However, although it appears to be reasonable to do so, the cases provide contradicting evidence. Company A and Company B both operate an LPS and have a digitalization strategy in place. However, in both cases, the digitalization strategy is not aligned with the lean strategy expressed by the company-specific LPS (C1-I.2) (C2-I.1). Company C has no dedicated digitalization strategy for its value creation processes. Consequently, no integrated lean – digitalization strategy exists currently.

The second question motivated by the aspiration for close interaction of lean and digitalization responsibilities is whether these individuals are linked from an organizational embedding perspective. Again, the assumption cannot be supported by the case results. All companies have a dedicated lean unit on a corporate level, Company A and B also have lean managers on a site level. The responsibility for the digitalization of the value creation processes is less simple to identify. While Company A and B perceive the teams responsible for introducing and optimizing the company-wide MES system as an internal driver of digitalization, in Company C the internal digital transformation is driven by the business processes automation team. However, all three companies have in common that lean and digitalization responsibilities are not organizationally linked (e.g., by a shared staff position) (C1-I.2), (C2-I.2), (C3-I.3).

The fact that lean and digitalization is neither considered holistically from a strategic nor from an organizational embedding perspective seems to be contradictory at first. However, as one case partner points out, it is quite reasonable as the tasks and core competencies of the lean unit are very different from those of the digitalization team. Another partner argues that an integrated consideration falls into the responsibility of the top management. If necessary, the top management delegates the responsibility to project teams. And indeed, temporary cooperation on a project basis is the preferred approach to ensure close interaction between lean and digitalization responsibilities. Company A brings together lean and digitalization specialists regularly when introducing a new process or new equipment (C1-I.2). Only after the new process was designed in accordance with lean principles was it digitalized by connecting the equipment and mirroring the process in the MES system. The same approach is found in Company B (C2-I.2) and C. The interview partner of Company C also points out that close cooperation between lean and digitalization specialists is strongly enhanced, if both teams understand the competencies of the other team and the potential value they can contribute to the project (C3-I.2).

5.4.2 Data Utilization in Manufacturing

This chapter comprises four subchapters. Chapter 5.4.2.1 consolidates the main drivers of data utilization in manufacturing. Chapter 5.4.2.2 summarizes the main challenges to exploiting manufacturing data, while chapter 5.4.2.3 discusses key enablers. Finally, chapter 5.4.2.4 discusses the results of the case studies on the impact of data utilization on LM.

5.4.2.1 Motivation

As discussed in chapter 4.1, a wide variety of DBAs exists and consequently, there is not one single objective but several objectives that might motivate the utilization of data in manufacturing. This chapter contrasts and summarizes the underlying motivation of the individual use cases discussed per case company. These use cases have been implemented only recently and are of high importance to the company, and thus, they are eligible to conclude the main motives of these companies to invest in data utilization. Besides these company-specific motivations, the experience of the two academic experts is used to shed light on the motivation for data utilization of a broader spectrum of companies.

5.4.2.1.1 USE CASE SPECIFIC MOTIVATION

Company A's primary motive to initiate the *Manufacturing Analytics Solution* project was to severely reduce the time and effort needed to access data for several analytics applications. Previously, the high effort for searching and combining relevant data made a majority of the analytics ideas uneconomical to realize. With the new system, employees are enabled and encouraged to perform their own analytics on an ad hoc basis without the need for long preparation times (C1-I.4).

The basic motivation of Company B's *Global Performance Cockpit* can be summarized as gaining supply chain transparency to minimize production stops. As the different production sites work in a network, some sites depend on the timely delivery of intermediate products of other sites. Hence, potential delivery problems need to be identified and communicated early. Therefore, Company B regularly assesses the OEE of critical equipment in all production sites and visualized the data on a cross plant performance cockpit (C2-I.4).

The main driver of the Smart Support system of Company C was to improve customer service and to free service employees' capacities. Therefore, Smart Support builds on historical data on customer problems and potential solutions to provide decision support to facilitate fast and sustainable solution suggestions to the customers. Enabling ad-hoc analytics (Company A), receiving early indicators for delivery problems (Company B), and providing decision support to employees (Company C) are three examples that demonstrate the motivation for using data is very diverse and highly dependent on the actual application.

5.4.2.1.2 EXPECTATION OF STAKEHOLDERS

To capture the motivation for data utilization of a larger set of manufacturing companies, this question was discussed with two recognized academic experts.

One main motive for DBA projects raised by both experts is rather trivial. Stakeholders of manufacturing companies (e.g., the supervisory board) increasingly pressure the top management to “do something in regards to Smart Manufacturing” (E2-I.1). Also, stakeholders increasingly expect from the top management to formulate a strategy on how to use manufacturing data (E1-I.1). As a consequence, initiatives for data utilization are regularly driven top-down. Often, however, as Dr. Schuster observes, the top management is rather keen to make sure that “something is done” than being actively involved in the data utilization initiative.

The manufacturing unit is usually responsible for identifying and implementing useful DBAs. Both experts agree that top-down driven initiatives for data utilization are often seen critically by the manufacturing units, and thus, an initial resistance is observed frequently (E1-I.2). However, this observation cannot be generalized as in other projects the manufacturing unit was welcoming of the initiative and immediately started to actively support the project (E2-I.2).

5.4.2.1.3 FURTHER DRIVERS

Besides top management's intention to fulfill external expectations as well as concrete improvement objectives, as seen in the company examples, three more drivers of data utilization in manufacturing have been identified. First, machinery manufacturers are equipping new machines with sensor technology as a prerequisite to offer additional services, such as remote and predictive maintenance, and new business models, such as pay-per-use (E2-I.3). Hence, the technical prerequisite for DBAs as well as "ready-for-use applications" for data utilization is pushed by suppliers into manufacturing companies.

The second argument to start investing in data utilization now are favorable economic conditions. Due to the significant reduction of cost for the required technological infrastructure, investing in data utilization has become attractive for a wide set of companies for the first time (E2-I.4).

The final driver is specific to companies in the USA and China. Both countries are leading in the field of ML (E2-I.5). Although the expertise is currently used by large internet companies and governmental agencies, it might be that in the long run the manufacturing industry in these countries may benefit from the know-how that is built around ML/AI.

5.4.2.2 Challenges

Following Hirsch-Kreinsen et al. (2018, p. 181), the identified challenges are grouped into three categories. The category *Employee* addresses all challenges that are related to an individual person, such as qualification and personal preferences. The category *Organization* comprises challenges that go beyond individuals but affect the whole organization. The category *Technology* includes technological requirements, such as the IT infrastructure.

An overview of the challenges is given in Table 31 in the consolidation chapter 6.

The individual challenges have been consolidated as far as possible into one over category. The over categories, in turn, were assigned to one of the three main categories Employees, Organization, and Technology. The assignment, however, is not exclusive. Employee qualification, for instance, can be considered both a challenge concerning employees and a challenge for the organization to organize the needed training to support employees in achieving the required level of qualification. Data security can even be discussed from all three perspectives, as it involves employee training to raise awareness (Employee), company guidelines for data protection (Organization), and technical infrastructure to avoid external intrusion into the IT system (Technology).

5.4.2.2.1 EMPLOYEES

C-E.1[30]. Initial Resistance of Employees

A significant challenge, which emerged during the qualitative studies, is ensuring employee acceptance of DBAs. The underlying assumption of the TAM model introduced in chapter 1.5.5 is that new technology is only used successfully if accepted by users. Both the company case studies, as well as the expert interviews, have revealed potential threats to employees' acceptance of DBAs. The initial resistance of employees may be caused by the fear of job loss, the fear of loss of autonomy in the job, and the fear of control during work.

C-E.1.1 Fear of Job Loss

As demonstrated by the *Smart Support* use case of Company B, DBAs have the potential to significantly increase the efficiency of an employee in terms of speed and required effort. Consequently, the assumption that the increased application of DBAs may result in a reduction of jobs is very understandable. Addressing these *fears of job loss* by demonstrating the value of DBAs (e.g., for competitiveness) is a key challenge for companies (C2-I.6). The senior researchers confirm the emergence of fears of job loss in the context of DBA, especially among shop floor employees (E1-I.2). For instance, maintenance employees may expect that predictive maintenance reduces the total amount of maintenance effort needed in the future and, as a consequence, a cut of maintenance jobs.

C-E.1.2 Fear of Loss of Job Autonomy

Fear of loss of job autonomy has also been documented by the Smart Support case. The support system was initially perceived as a threat to the autonomy of service employees. The perspective of being downgraded from a competent, self-determined solution developer to a mere interface to communicate solutions proposed by the system to the customer, was not welcomed by the employees *(C3-I.5)*. Similarly, a DBA use case of predictive maintenance of Company A was rejected by maintenance employees. As noted by the interview partner, maintenance employees did not reject the new DBA due to technological inabilities, but because they disliked not being involved in the development of the new approach (p.129).

Even some manufacturing managers feel uncomfortable with the concept of data analytics as a basis for decision-making since the request for using data-based decision support systems is perceived as questioning their competence (E1-I.2).

---

30 Each challenge has a unique identifier to allow linking enablers to the corresponding challenge.
C-E.1 stands for first main challenge of the category employee, C-E.1.1 is referencing to the first sub-challenge of the first main challenge of the category employee.

C-E.1.3 Fear of Control

The third challenge in this category is *fear of control*. The potential of misusing manufacturing data to monitor the behavior of individuals on the shop floor is as a general barrier for data utilization in manufacturing (C1-I.9). As a representative of Company A observes, employees are welcoming digital technologies and data analytics in general, but the acceptance is very limited if data can be used to monitor individual behavior and performance. The challenge is linked to the challenge *Misuse of data* which is discussed in the section on organizational challenges.

C-E.2. Employee Qualification

Employee Qualification was raised by all three case companies as well as the two experts as a key challenge for companies striving to implement DBA successfully. The following section discusses the qualification requirements of three different groups of employees: shopfloor employees (1), who are the end-users of DBAs; managers from middle management (2), who usually drive the implementation of DBA projects; and top managers (3), who often initiate and sponsor the projects.

C-E.2.1 Shopfloor employees

New technologies and applications require additional qualifications for end-users. Thereby, an adequate qualification is necessary for at least three reasons: first, end-users need to be trained to use the new application effectively and to exploit its full potential (C1-I.12). For instance, employees of Company A require basic training on the functionalities of the new MAS system to be able to perform analysis on their own.

Second, as data quality is crucial for the quality of the result of DBAs, employees need to understand the importance of entering correct data into the system. By gaining a rough understanding of how a DBA is processing the manually entered data and thus to be able to link entered data to outcome results, employees can better understand the implications of their work. This understanding significantly increases the motivation to invest effort to input manual data correctly (C2-I.10).

Finally, shop floor employees are a valuable source of information for DBA developers, as they have the best understanding of the actual conditions and pain points on the shop floor. However, only by understanding the underlying logic of a DBA, shop floor employees are able to provide qualified feedback and change requests to developers (C1-I.12).

C-E.2.2 Middle Management

While DBA projects are often initiated by the top management, it is usually the task of the middle management to drive the actual implementation (p. 156). Due to the interdisciplinary nature of DBAs, driving DBA implementation projects requires a broad understanding, not only of the manufacturing system and the IT system but also of data analytics methods (C2-I.11).

In order to overcome initial resistance of employees (see above), change management is an additional skill that appears to be of high value for middle managers driving DBA projects. Although middle managers do not need to be an expert in every field discussed above, they need to have a good understanding to set the right directions. Identify the right individuals from the middle management to drive DBA projects and provide sufficient training to meet the diverse requirements constitutes a challenge for many companies. As Company C reports, the demand for advanced training in data analytics is currently exceeding the offer by far (C3-I.9).

C-E.2.3 Top Management

A key challenge of many companies striving for DBAs is a lack of technological expertise on higher management levels. The missing technological expertise results in the inability of top managers to critically evaluate the usefulness of a DBA project (p.153). As a consequence, Professor Schuster perceives the decision for or against a DBA project often as “a question of faith.” If they believe the consultant, managers tend to follow his advice uncritically.

5.4.2.2.2 ORGANIZATION

C-O.1. ROI Calculation of Investment Decisions

A key barrier for DBAs, which is identified by both academic expert interview partners consistently, is the uncertainty of the ROI of DBA projects. This is mainly due to two reasons: first, due to the uncertainty of results and second, due to the chicken and egg problem that applies to many DBAs.

C-O.1.1 Uncertainty of Results

In traditional investment decision-making, managers used different approaches to forecast the value that is likely to emerge as a result of a specific project. By translating the expected value into financial numbers and contrasting investment costs and expected financial benefits, managers derive the decision in favor of or against a project proposal. This basic concept reaches some limitations in the case of data analytics DBA projects, as their outcomes, cannot be predicted seriously in advance (E2-I.7). As a consequence, calculating the ROI and forecasting the expected break-even point of investments into DBA projects is afflicted with a high level of uncertainty. Therefore, the risk of such investments and the possibility for unprofitable investments is higher than in projects that do not rely on data analytics findings.

C-O.1.2 Chicken and Egg Problem

The uncertainty of results leads to a second challenge that is inherent to DBA projects comprising data analytics, which can be summarized as a chicken and egg problem (E1-I.10). As company managers usually like to have an estimate of the ROI and a clear business case before granting project funds, they are very reluctant to do so if no ROI can be presented, or if the ROI calculation is afflicted with high risks. However, consultants or other individuals that intend to implement the DBA are only able to demonstrate the value of the application after they had access to the data and found useful patterns in the data. From a theoretical point of view, this situation prevents investments in DBA projects that are not only for the purpose of learning but have to compete for scarce resources with other investment proposals.

C-O.2. Misuse of Data

C-O.2.1 Internally

Tracking more and more data from the manufacturing system facilitates transparency on the performance of the production system. However, the same data might be used to monitor the performance not only of the system but also of individual employees (C3-I.7). Superiors that are suspicious that some of their workers are underperforming may feel tempted to use this new opportunity. Furthermore, data might be misused to create a relatively accurate picture of the individual behavior of a shop floor worker (C1-I.9). Misusing data internally for employee control not only lowers the motivation among workers but can also have legal and business implications. In countries with strict data privacy regulations, companies can be sued for misuse of personnel-related data. Furthermore, at least in the EU, labor unions are powerful enough to block even major projects if they do not feel privacy rights are protected sufficiently (p.137).

C-O.2.2 Externally

Protecting data from being accessed from unauthorized persons from outside the organization is a central theme and is perceived as one of the greatest challenges for data utilization in manufacturing by Company A (p.127) and B (C2-I.9). Although data security is mostly discussed from a technical point of view, our case partners emphasize that often the weakest part in the firewall is not the IT system but employees. By methods like social phishing, employees are tricked into passing on access authorizations or data to unauthorized persons. Preventing data theft from external actors by creating awareness of the risks of data theft and providing employees with guidelines is at least as critical for data security than keeping the IT system up to date.

C-O.3. Comparability of Data

Several DBAs require combining data from different data sources. Hence, the challenge arises to make datasets from different sources comparable by standardization.

C-O.3.1 Within the Plant

An example of the required standardization of data referred to in the case studies are failure codes (p.125). The integration of failure data of several machines of the same kind in order to enhance the effectiveness of a data-based root cause analysis requires fully standardized failure codes.

C-O.3.2 Within the Manufacturing Network

As seen in the data utilization use cases of Company A and B, data may not only collected and combined within the borders of a factory. The *Global Performance Cockpit* allows monitoring performance metrics from different production sites within the manufacturing network. To ensure comparability of metrics from different sites, those metrics need to be standardized not only within one factory but within the whole manufacturing network (C2-I.5).

C-O.4. Access to Expert Knowledge

C-O.4.1 Limited Internal Resources

Developing DBAs requires specific expertise in different fields of data analytics and IT, depending on the actual use case. These skills can hardly be trained in a few weeks but rather require a full-time study. As the field of data analytics is comparably new, many companies face the challenge to have insufficient internal data analytic experts. Even Company C, which is working with data for many years, struggles to find and train sufficient data experts internally. Our interview partner put it concisely: “We know that we need more data-affine employees” (p. 146).

C-O.4.2 Strong Competition for Data Scientists

The limited internal resources of data experts, the high demand, and the fact that professional training of data scientists is very time demanding suggests the approach of hiring new data experts from the labor market. As shown in a study by Macuvele et al. (2018, p. 39), almost every second company is already hiring employees with data analytics skills or intends to do so. However, professional data scientists are not only wanted by manufacturing companies but also by financial institutions and consultancies. The high demand for data scientists and the scarce availability on the labor market results in a “war for data talents.” Consequently, hiring data scientists is currently very challenging and expensive due to high salaries paid by other employers. This observation was made by the manufacturing company A (p.130), as well as the ICT company C (C3-I.10).

5.4.2.2.3 TECHNOLOGY

C-T.1 Basic Requirements

Within the qualitative studies, no surprising findings regarding the technological challenges have emerged. On the contrary, the main technological challenges raised by the interview partners are in line with the challenges identified in the literature, based on the individual DBA use cases (see chapter 4.1.6) (E1-I.7).

However, two technological challenges are emphasized as essential basic requirements for data utilization: IT system performance and data security.

C-T.1.1 IT System Performance

In this context, IT system performance addresses the ability of the IT system to transfer, store, and process large amounts of data within a reasonable amount of time. The definition of a reasonable amount of time is highly dependent on the use case of data utilization. The performance of the IT system is mainly determined by the bandwidth of the data connection and by the computing power. Hence, very complex DBAs, for instance, complex simulation problems, as well as DBAs relying on near real-time data require high performing IT systems (E1.I-6).

C-T.1.2 Data Security

As discussed in chapter 4.1.6, the literature considers data security as a critical challenge for utilizing manufacturing data (Ghobakhloo, 2018, p. 921; Sommer, 2015, p. 1515; Thoben et al., 2017, p. 12). This observation is in line with the perception of all interview partners, that data security is a key technological challenge (C2-I.9). As the case and interview partner are experts in the area of lean and manufacturing rather than IT, they did not go into detail about how to ensure data security from a technological perspective. However, to complement the data security efforts of the IT department, employees are briefed to be aware of the risks of data theft (e.g., by guidelines to avoid social engineering or phishing). Access and identify security are included in mandatory training for all employees (p. 173).

C-T.2. Distributed Data

The rationale for the introduction of the *Manufacturing Analytics Solution* system of Company A was the enormous effort required to locate and merge distributed manufacturing data (C1-I.4). A majority of the ideas for data analytics were not realized as the initial effort to create the necessary database was not economically justifiable. This data utilization use case sheds light on two challenges companies need to address to minimize the effort for analyzing distributed data. First, integrating data from different sources and databases (machines, IT systems, production sites) and second, to enable the employee to access the data relevant for their job.

C-T.2.1 Data Integration

Integrating data from several data sources and IT systems was a key challenge of Company B during the implementation of the *Global Performance Cockpit* (C2-I.5). Thereby, defining interfaces between different IT systems, or combining several systems into a new one, has shown to be complex and time-consuming.

A particular form of data integration, more specific knowledge integration, was needed to realize the *Smart Support* use case of Company C. Historically, the knowledge about customer complaints and solutions was distributed among the service employees in the form of implicit human knowledge. Translating this implicit knowledge into explicit machine-readable knowledge that can be used for ML application was a key technological challenge for the *Smart Support* case (C3-I.4).

C-T.2.2 Data Access

The second challenge to facilitate manufacturing data analytics is to create a fast and user-friendly interface for employees to get access to the data relevant for their intended analysis (p.127). Besides providing the technical opportunity to access a central database remotely, companies need to balance the desire to access as much data as possible and the risk of data theft or misuse if too many people have access.

C-T.3. Miscellaneous

The two senior academics serving as expert interview partners have gained in-depth knowledge and a broad range of experience in their field of study. Due to their involvement in several industry projects, they are capable of recognizing specific challenges of data utilization in manufacturing, which are not evident at first. The last category *Miscellaneous* lists two of these challenges, first the problem of insufficient failure data for a predictive maintenance application and second the challenge to find algorithms that are applicable in the context of the manufacturing industry.

C-T.3.1 Insufficient Data

Predictive maintenance relies on a large set of machine error data to identify patterns and to forecast the optimal time for maintenance. However, if the availability of a machine is of high importance, companies tend to invest ample maintenance effort to keep the machine running. Consequently, this machine is likely to generate none or only very few machine failure data. This has several negative implications for the predictive maintenance application (E1.I-5).

First, patterns identified on the base of a small database are less robust. Second, with few failure data, it is challenging to assess whether the predictions made by the predictive maintenance application are correct or not. As a consequence, nobody can seriously assess if the low number of machine failures is a result of a well-working predictive maintenance application or rather the result of much or even unnecessary much maintenance effort. In that case, predictive maintenance is not able to fulfill a central value proposition, which is to determine the optimal balance between investing too few and too much maintenance efforts.

C-T.3.2 Inapt Algorithms

A majority of state-of-the-art ML algorithms are developed for contexts different from the manufacturing industry. ML algorithms developed from internet companies are designed for large samples and low complexity. As discussed above, the sample size of data in manufacturing can be comparably small, while the structure of the data is often more complex. Developing tailored algorithms will, according to Professor Wuest, pose a challenge to the whole manufacturing industry.

5.4.2.3 Enablers

According to the Cambridge University Press (2019b) dictionary, the term enabler is defined as "something or someone that makes it possible for a particular thing to happen or be done."

In this context, an enabler is a measure that increases the likelihood of the successful implementation of DBAs. In the previous chapter, challenges for data utilization that have emerged in the qualitative studies have been presented and discussed. This chapter builds on this work and consolidates enablers that are useful to address a large part of the identified challenges. Thereby this chapter follows the same basic structure by grouping the enablers in the three established categories Employees, Organization, and Technology.

Like the challenges in chapter 5.4.2.2, the enablers in this chapter are derived directly from the qualitative studies. The enablers are discussed in more detail below. An overview of the enablers, including a link to the corresponding challenge, is presented in Table 32 in the consolidation chapter 6.

5.4.2.3.1 EMPLOYEES

E.1. Foster Acceptance

The cross-case analysis of the three case studies and the two expert interviews has uncovered three threats to employees' acceptance of new DBAs. Following the TAM proposition that the efficient use of technology depends on the users' attitude to use the system, companies need to foster the acceptance of new DBAs by addressing the three subchallenges of the main challenge *Initial Resistance of Employees*. The sub-challenge *Fear of Job Loss* might be addressed by communicating the benefits and the added value of new DBAs. Integrating end-users in the development of a new DBA solution decreases the *Fear of Loss of Job Autonomy.*

The third sub-challenge *Fear of Control* is covered in the Organization section as part of data transparency and utilization guidelines.

E.1.1 Understand the Value of DBAs

To address the *Fear of Job Loss,* it is a key leadership task to communicate the benefits of a new DBA for the whole company as well as for every employee. By demonstrating how a new DBA contributes to the overall company's competitiveness, the *Fear of Job Loss* can be mitigated. An effective change management that is "tak[ing] the people on board" (Company B) is even more important, if a new application requires extra effort by shop floor employees (e.g., to enter data manually) while the benefits for that effort are hard to see from an individual point of view (p.136).

E.1.2 Being Part of the Solution Development

A second cause of lacking employees' acceptance for a DBA can be the perception that the individual role and autonomy for finding the best solution is reduced due to data-based decision support systems. Company A (Predictive Maintenance application) and Company C (Smart Support system) have witnessed rejection of DBAs due to this reason in the past.

A valid approach to mitigate the fear of employees of being deprived of their influence is to allow an active involvement in the development of the DBA of the final end-users. Following the example of Company A, maintenance staff is more likely to welcome predictive maintenance solutions if they have been personally involved in the development compared to a solution that was introduced top-down without the opportunity for feedback. As different restrictions do not allow to include every shop floor employee in the design and development phase of DBAs, the opportunity to provide feedback is fundamental: “introducing a new solution without collecting feedback first, does not work” (C1-I.8). Hence, training and feedback workshops are organized regularly when a new solution is introduced.

Company C observed resistance of service employees when introducing the Smart Support system as some feared that their role will change from determining the best solution for the customer to just communicating the output of the new system. Taking these concerns into account, Company C gave its service employees the freedom to decide whether to follow the system's proposal or to propose an individual solution based on personal experience. The approach resembles the concept of an autonomous driving car, which allows the driver to take over control whenever he or she feels it is necessary. Hence, employees do not feel their job autonomy diminished, although they follow the system's recommendation most of the time. By now the Smart Support system is perceived as a useful support tool and is generally accepted among service employees (C3-I.5).

E.2. Role-specific Training

Analyzing the challenge to adequately qualify a company's employees has revealed different requirements for different hierarchical levels. This section proposes an approach of role-specific training, tailored to the individual needs of shop floor employees, middle managers, and members of the top management level. However, equally relevant for all employees who deal with data is training to raise awareness of the risk of data and or access credential theft and to provide guidelines for individual behavior to protect data and access credential.

E.2.1 Basic Training of End-users

Employees require additional qualifications to work effectively with new technologies (E2-I.11). Furthermore, understanding the basic logic behind a DBA allows end-users to use the application more effectively as well as to provide more qualified feedback to the solution developer (C1-I.12). This not only improves the user-friendliness of a new DBA but also increases its acceptance among end-users. Accordingly, companies should not only train employees in using an application but also to understand their underlying logic (C3-I.8). Advanced training for end-users (e.g., to fully understand the ML foundation of the Smart Support tool) does yield little additional value for the daily job and is therefore economically not viable.

E.2.2 Citizen Data Scientists

A promising approach to built-in data analytics skills has emerged in the qualitative studies and is referred to as *Citizen Data Scientists*.

The basic concept is simple. Very motivated employees who already have prior knowledge in IT or data analytics receive extra training in data analytics. Afterward, they serve as internal advisors for other employees, who plan to implement a data analytic application. The approach was suggested by academia and industry representatives. Professor Wuest argued from a theoretical point of view that technicians and engineers already have a good mathematical understanding and thus are well-prepared for additional training in data analytics (E2-I.12). He points out that training selected technicians intensively in data analytics allows companies to combine existing manufacturing know-how with data analytic skills. In this context, he suggests to offer motivated employees a range of voluntary training programs, thus allowing the employee to determine their *optimal* level of qualification for their tasks individually (E2-I.11).

Company A has already introduced the concept of Citizen Data Scientists in one of its lead factories for digitalization. The word *Citizen* refers to the fact that they are not professional data scientists with a formal education in data science. Basic requirements to receive the additional training are a strong intrinsic motivation, high IT-affinity as well as basic IT-skills. Selected employees receive extensive on-the-job training in data-analytics tools but are also requested to perform self-organized training to keep their knowledge up to date. Company A supports Citizen Data Scientists by providing additional resources to the employee, such as extra time dedicated to training and, if required, a budget to also select external, fee-based educational programs. In addition to courses and self-learning, regular exchange within the plant and with data analytics experts from central company departments is part of the training (C1-I-7).

When reached a certain level of expertise, Citizen Data Scientists serve as internal consultants and sparring partners to all topics related to data. Internally, they are referred to as “the masters of data.” Citizen Data Scientists thereby take an intermediate role between end-user and middle-manager. On one hand, they have a deep understanding of the needs of shop floor employees and will act as end-users of DBAs themselves. On the other hand, they support middle managers in driving the implementation of DBA projects by combining manufacturing domain and data analytics expertise.

This approach has supported Company A to overcome the challenge of a lack of data analytics expertise within its lead factory for digitalization. Due to their expertise in manufacturing and data analytics, Citizen Data Scientists enjoy high appreciation and, as a result, are often offered jobs from central units and competitors (p.130).

E.2.3 Advanced and Holistic Understanding at the Middle Management Level

From his experience of several data analytics projects with partners from the manufacturing industry, Professor Schuster has presented some characteristics that distinguish companies that are successful in implementing DBAs from companies that fail to do so.

Accordingly, a key enabler of a successful DBA project is the availability of relevant technical expertise within the company. As top managers often lack this expertise as well as the time, the responsibility to drive DBA projects is usually passed on to the middle management. Consequently, it is critical to have managers on this level with an advanced understanding of data analytics. The availability of expertise on this level is crucial for two reasons: first, to allow managers to serve as critical sparring partners for consultants or service providers and second, to drive DBA projects internally. However, as shown in the challenge *C-E.2.2 Middle Management,* DBA projects are at the interface of manufacturing, IT but also require change management skills to overcome initial resistance. Thus, the ideal driver of DBA projects combines manufacturing and IT know-how, leadership qualities, and as a new requirement, data analytics skills.

As leaders with data analytics skills are still rare, Company C has created a two-pillar training system for managers from the middle management level. The first pillar comprises formal training. In cooperation with a well-known university, a specific 150-hour course was developed to train employees in data science. The second pillar consists of building competencies internally by fostering internal knowledge exchange. For instance, interested employees can choose to work for a maximum of 40 weeks for a fixed percentage of their time in an IT or data science team. The training offer is highly appreciated by employees, resulting in a large request for participation in the university course that is currently exceeding the more than 100-course places by far (C3-I.9).

E.2.4 Basic but Holistic Understanding at the Top Management Level

The lack of technical expertise of top managers has been discussed in challenge *C-E.2.3 Top Management.* However, following Professor Schuster, there are no simple solutions to this challenge. A full content course, such as the course presented in the section above, would be challenging to fit into the calendar of most top managers. Besides, such a course requires significant existing knowledge in the fields of mathematics and software. For that reason, he expects few participants from the top management level in such courses.

However, a course with a reduced scope, tailored to the special requirements of top managers is expected to be valuable and attractive for many companies. Even by understanding only the basic concepts of data analytics, managers will be better prepared to evaluate internal and external DBA project pitches adequately (E1-I.11). Ideally, additional expertise can be consulted internally if required.

5.4.2.3.2 ORGANIZATION

O.1. Favorable Conditions

As discussed in challenge *C-O.1.1* *Uncertainty of Results* the outcomes of data analytics DBAs are not predictable and desired outcomes cannot be guaranteed. There is no strategy to avoid the risk of finding no useful patterns in the data.

However, DBA projects can fail for several other reasons, too. Thus, creating favorable conditions by mitigating as many potentials barriers as possible reduces the risk of failure of DBA projects. This section consolidates four enablers from an organizational perspective to create favorable conditions for DBA projects. First, applying a structured approach to select potential DBA projects. Second, to ensure management buy-in—not only to start a project but also to overcome difficult project phases. Third, having the right kind of leader in place. And fourth, maintain the motivation in case of setbacks.

O.1.1 Structured Approach to Select DBAs

As seen in chapter 4.1 and the three data utilization use cases, a wide variety of potential DBAs exists. Selecting the most promising DBAs is an essential first step for implementing DBAs successfully. Thereby, a structured process can help to identify potential DBAs and to evaluate their economic rationale. In case of Company A's digitalization lead factory, this process is called *Analytic Business Case Review* (C1-I-10). As part of this review process, the plant responsible for digitalization (including data analytics) gets in contact with value stream managers in the plant. At a joint meeting, he presents successful practice examples of data analytics solutions that have already been tested and rolled out successfully in other areas. Acting as advisor and sparring partner, he discusses the potential of the set of successful practice use cases with the respective value stream manager. Based on experience, the digitalization responsible has a good sense for the required human and financial effort, thus supporting a realistic cost-benefit consideration. Only if both managers recognize enough potential to justify a new project, the process goes on.

In a second step, the project is defined in more detail, and if required, additional resources are requested from the plant leadership team. During the implementation phase, the project is then supported by the internal Citizen Data Scientists. Besides a well-informed discussion on the basis of existing use cases and a realistic cost-benefit consideration, a third advantage of this process was expressed. By including the value stream manager early in the *Analytic Business Case Review* process, the decision on the implantation of a new DBA is supported at the operational level.

O.1.2 Ensure Management Buy-in

DBA projects often require the investment of initial effort to collect the necessary data, before any analytics can be performed. However, the uncertainty of data analytic outcomes results in little willingness of the management to grant a budget to cover for the initial effort. To overcome the chicken and egg problem (see challenge C-O.1.2), Professor Schuster proposes to provide success practice examples of similar situations and is thus in line with the *Analytic Business Case Reviews* approach. Providing a proof of concept facilitates managers to overcome the dilemma of investing resources in projects with uncertain ROIs.

Top management support is still critical after the initial phase for at least two reasons. Some DBA projects require the adjustment or new implementation of IT systems (e.g., Global Performance Cockpit case of Company B), which is often very time-consuming. Second, due to the uncertainty of DBA results, management commitment is inevitable, especially if DBA projects do not deliver the expected results in the first attempt.

As discussed in chapter 2.5, data analytics is not a linear but an iterative process (see Figure 7:) and finding useful results may require more than one iteration round. Therefore, a lack of management commitment can result in a hasty termination of the project. According to Professor Schuster, companies with a high persistence against setbacks in the early phase of DBA projects are more likely to implement DBAs successfully.

To foster continuous management support, Professor Wuest proposes to build in “Quick Wins.” This approach will be discussed in enabler *O.1.4 Maintain the Motivation.*

O.1.3 Put the Right Leader in Place

The selection of the leader in charge of a DBA project strongly influences the project's probability of success. Summarizing his personal experience from joint DBA projects with industry, Professor Schuster presents a profile of requirements for a good DBA project leader.

First, DBA projects are too time-consuming to be driven in addition to the daily business. Therefore, a DBA project leader should have dedicated time to manage the project. Second, a certain level of seniority of the project leader does certainly help to have access to enough financial and personal resources. Furthermore, seniority is valuable to convince the top management for support as well as to overcome internal resistance against DBA. Third, persistence (“a certain terrier mentality”) and persuasiveness is a necessary character trait of a leader, not only to overcome initial resistance but to keep the project running despite potentials setbacks. (E1-I.8).

Professor Wuest agrees and especially highlights the two key characteristics of persistence and access to sufficient human and financial resources as decisive (E2-I.9).

Again, based on personal experience as a consultant for DBA projects, Professor Schuster notes that members of the middle managers who meet the requirements described above and have deep technical expertise have been the best project partner so far. In contrast to representatives of the higher management, middle managers may not only meet the key requirements of a good DBA project leader but are usually rather in the position to invest the time needed in the project.

O.1.4 Maintain the Motivation

Because DBA may not deliver the expected results within the planned timeframe, the motivation of project members and other stakeholders can diminish. However, as highlighted by both expert interview partners, persistence is an essential feature to implement DBAs effectively. According to Professors Schuster and Wuest, little success stories along the way, so-called *Quick Wins*, mitigate the risk of falling motivation, even if the desired outcome is not (yet) reached. “‘Quick Wins keep people happy and management calm” (Professor Wuest).

However, natural Quick Wins are rare (p. 153) and therefore need to be built into the DBA project from the start (E2-I.10). An example of a Quick Win presented by Professor Wuest is a comparably simple threshold-based control system, which is developed as part of a more complex predictive maintenance DBA. In case the DBA does not achieve to fulfill the requirements instantly, the project is not considered as failure, as some additional value is created by the threshold-based control system. Arguably, the motivation to further invest the effort to achieve the full potential of the DBA is higher with small success stories along the way than without any reward on the previous investments.

O.2. Data Transparency

The challenge *C-E.1.3 Fear of Control* has demonstrated the risk of intransparency of data utilization in manufacturing to the acceptance of employees of DBAs. During the qualitative studies, two approaches have briefly discussed to mitigate the risk of data misuse. Binding data guidelines and establishing the role of a Data Security Officer.

O.2.1 Data Guidelines

To avoid misuse of personal data, Company A has created data guidelines. The guidelines are supposed to create transparency over the collection and utilization of data. Furthermore, the guidelines define clear limits for tracking individual behavior and performance. Creating transparency includes open discussions with employee representations, with the results of the discussion informing the guidelines. The strict and transparent data policy supported the plant management in ensuring acceptance among its employees for DBAs.

The fact that growing data collection is increasing the transparency of individual performance is not denied by Company C. However, transparency is also perceived as a chance to establish a fair and transparent performance evaluation system. Transparency and guidelines on the data collected and the performance metrics used for individual performance assessment is key to secure employee acceptance for that approach (p. 145).

O.2.2 Data Security Officer

In addition to establishing data guidelines, each plant of Company A has its own Data Security Officer (C1-I-11). He or she is in charge to ensure full compliance with internal and external data security and privacy regulations.

O.3.Standardization

As shown in challenge *C-O.3 Comparability of Data* many DBAs require standardized and comparable data, for instance, failure codes and performance metrics.

O.3.1 Management Responsibility for Standardization

Standardization of metrics company-wide may require some production sites to change their definitions and calculations. Usually, sites are hesitant to change internal definitions as it takes much effort to modify all calculation processes and also because the new metrics are not compatible with legacy data anymore. Consequently, setting and enforcing company-wide standards need to be driven and ensured by the management on a corporate level (C1-I.6).

O.4. Internalize external Expertise

The challenge *C-O.4. Access to Expert Knowledge* has illustrated the difficulty of the case companies to meet the internal demand for employees trained in data analytics skills. The approach of training their own employees has already been discussed in enabler *E.2. Role-specific Training.*

The second option is to internalize external expertise. However, strong competition for data scientists on the labor market makes it challenging for manufacturing companies to hire as many externals as desired, especially due to solvent competitors from other industries. To convince potential candidates despite financially more attractive job offers, manufacturing companies need to highlight non-monetary benefits. Alternatively, companies can benefit from external expertise by collaborating with external partners.

O.4.1 Convince with non-monetary Benefits

Due to collective agreements, Company A is somehow restricted to offer very high salaries, which are out of the standard range. Financial institutions and consultants, in contrast, can offer salaries that exceed the maximum limit of those of Company A. Nevertheless, Company A manages it to attract and hire highly qualified data scientists. A key argument for some candidates is the fact that they can see the result of their work come to a reality in the plant "within 100 meters of their working place." The chance to test new ideas quickly in a physical setting and to receive real-world feedback instantly is a strong argument in favor of Company A (p.183).

Also, Company A and Company C have observed that a high reputation as an employer is beneficial for hiring highly demanded specialists.

O.4.2 Benefit from Cooperation

Professor Wuest recommends companies to partner up with external partners in order to benefit from their experience and resources. Companies can collaborate, for instance, with consultancy companies and research institutions. Maybe biased by his role, Professor Wuest argues that universities are attractive collaboration partners for manufacturing companies. Master and Ph.D. students not only bring a high intrinsic motivation to complete a project successfully in time but also have in contrast to company employees the time to work on it almost full time.

Furthermore, universities may have access to experienced experts in the field of data analytics as well as software and computing power. As their primary objective is progress in academic research, universities are eager to participate in projects not primarily for financial reasons and are thus often more affordable as cooperation partners than consulting companies.

5.4.2.3.3 TECHNOLOGY

From the six technological challenges (see chapter 5.4.2.2.3), only for the challenge *C- T.1.1 IT System Performance* a corresponding enabler was discussed in the case studies. Cloud computing is a promising technology to provide companies access to scalable computing and storage capacity. An external enabler for DBAs is the significant reduction of costs for technical equipment in the last 10 years.

T.1. State-of-the-art Technology

T.1.1 Benefit from falling Component Costs

The technical infrastructure to collect, transfer, process, store, and protect data (see chapter 4.1.6) remains a key prerequisite for DBAs. However, as Professor Wuest notes, referring to the final report of the World Manufacturing Forum 2018 (Taisch et al., 2018, p. 15), prices for technical equipment have decreased significantly in the last decade. Sensor prices have been halved in the last 10 years. The costs for data transmission bandwidth were reduced by a factor of 40, the costs for data storage by a factor of 50, and the costs for computer performance for data processing by a factor of 60. These significant cost reductions make these technologies affordable for more and more companies and are, thus, an external enabler for DBAs (E2-I.4).

T.1.2 Cloud Computing

Cloud computing was introduced in chapter 2.3.3 as a core technology of SM. Indeed, the case studies have shown that cloud computing can serve as a technological enabler for DBAs.

One example from an industry partner and one example from an academic partner underline the increasing relevance of scalable IT resources. Scalable computing and storage capacity have been key enablers of the MAS system of Company A (C1-I.5). Following Professor Schuster, scalable computing resources will also be increasingly relevant for simulation applications as their complexity increases exponentially when adding new elements (E1-I.6).

To cover the increasing and also fluctuating demand for computing and storage capacity, fixed IT resources are not flexible enough and either not powerful enough or too expensive. Cloud computing services will enable companies to use and pay exactly the amount of computing power and storage capacity needed for their DBAs.

5.4.2.4 Impact on Lean Manufacturing

This chapter consolidates the results of the case studies on the impact of data utilization on LM. Therefore, this chapter comprises two parts. The first part presents a summary of the general expectation of the case companies, how data utilization will influence LM. The second part then summarizes the impact of 10 real-world DBAs on different lean elements. The term *lean elements* comprises lean principles (see chapter 2.2.1) and lean practices (see chapter 2.2.3).

5.4.2.4.1 GENERAL EXPECTATIONS OF CASE COMPANIES

Lean remains lean. Company A expects no significant changes to the basic principles of LM. Lean practices that support lean principles, however, are expected to benefit from data utilization in manufacturing. This statement is supported by the findings exhibit in Table 28 in part two of this chapter. It provides four real-world examples of DBAs supporting lean practices from one plant of Company A. The single most important impact of data utilization on LM is increased transparency. The new MAS system, for instance, will allow Company A to perform data analysis more efficiently. Hitherto, it was challenging to evaluate the effectiveness of different measures of improvement. Accurate data from the MAS system, in combination with advanced analytic tools, allow this evaluation to be more reliable and more efficient. Also, the data will increase the transparency on inventory and thus facilitates the identification of opportunities for inventory reduction.

Full transparency is also the main value proposition of data utilization regarding lean for Company B (C2-I.12). Near real-time data analytics performed on accurate data will support the lean continuous improvement process.

“Gaining direct data-based insights on where the biggest potentials for optimization and the next improvements are, as opposed to relying on opinions and prioritizing them,” is the expected main benefit of digital technologies on LM. Accordingly, Company C expects a strong positive impact of data utilization on the identification of improvement potentials (C3-I.11).

Besides the use cases of DBAs discussed in the case study in detail, Company C applies several advanced DBAs to support the lean objectives of customer value and waste reduction. Among these are an intelligent chatbot to reduce waiting times of customers for written requests and a voice recognition ML application to reduce customers' effort to identify themselves on the telephone.

5.4.2.4.2 IMPACT OF DBAS ON LEAN ELEMENTS

The second part of this chapter summarizes use cases of DBAs documented in the case studies with a special focus on the impact of lean elements.

Table 28 provides an overview of nine use cases of data utilization, the corresponding DBA according to the classification in 4.1.3, the impacted lean element, and a brief description of the impact.

Table 28: Impact of DBAs on lean elements (documented in the case studies)

C	Application Name	Corresponding DBA	Impacted Lean Element	Impact	P
A	Manufacturing Analytic Solution system (MAS)	System Performance Measurement	Waste reduction, Continuous Improvement	Increased transparency allows a better identification of waste and supports CI by allowing to asses the effectiveness of CI measures.	131
A	Predictive Maintenance	Predictive Maintenance	Preventive Maintenance	Increased accuracy of maintenance planning.	131
A	Advanced Planning and Scheduling	Production Planning	Pull/Kanban	Pull is replaced partially due to the central production planning of critical parts.	131
A	E-Kanban	Material Flow Management	Pull/Kanban	Support of Pull/Kanban due to the fast and secure transfer of Kanban signals and the ability for dynamic lot size adaptions.	131
B	Global Performance Cockpit	System Performance Measurement	Continuous Improvement	Full transparency of data and processes on the shop floor supports the lean continuous improvement process.	138
B	Predictive Maintenance	Predictive Maintenance	Continuous Flow	Predictive Maintenance increases equipment stability which in turns enables Continuous Flow.	139
B	Production Planning System	Production Planning	Pull/Kanban	Pull is replaced between process steps due to central planning of the material flow.	139
C	Smart Support system	Product Quality Improvement	Create Customer Value	Based on historical quality data, a ML algorithm support the identification of the best solution, thus increasing customer value.	143
C	Performance measurement of non-physical processes	System Performance Measurement	Continuous Improvement	Comparable to shop floor management KPIs, the performance of non-physical processes is measured and visualized.	147

C: Company, P: Reference to the respective page

5.4.3 Commonalities and Differences: ICT vs. Manufacturing Companies

Case Company C is an ICT company and although they consider the physical infrastructure to provide its services as their manufacturing system, the value creation processes are mostly non-physical. This chapter contrasts the findings of the three case studies to identify commonalities and differences between Company C and the two manufacturing companies A and B.

Commonalities

First, all three companies have reported to advocate a close interaction between digitalization specialists and lean specialists. The benefit of making a process first efficient and robust with lean approaches before digitalizing it is stated consistently. However, none of the three companies has formalized an integrated perspective on lean and digitalization, neither from a strategic perspective nor from an organizational embedding perspective. All three have a central lean unit on the corporate level, while the responsibility for digitalization cannot be clearly assigned to a unit. Interaction between lean and digitalization responsible is fostered and realized on a project basis in all three companies. Mostly, these project teams are established when a new tool or process is introduced.

Second, the perceived threat of loss of job autonomy is shared between employees of Company A and Company C. Service employees of Company C feared that their role will be impacted negatively due to the introduction of the Smart Support system. Similarly, maintenance employees of Company A rejected a data-based predictive maintenance tool, which was introduced top-down but preferred instead to rely on their own experience. The response of both companies to the initial resistance is also similar, as both ensured their employees to be part of the solution. The Smart Support users always have the power to overrule the suggestion by the system. Company A reacted to the situation by involving stakeholders in the development process of new applications, thus allowing them to provide earlier feedback. Being involved in the solution development increases the acceptance of an application, regardless of whether in manufacturing or in customer service.

Third, the request for role-specific training is present in both companies. End-users should receive basic training that includes the understanding of the underlying logic of a new tool. Thus, end-users are not only able to use the tool more effectively but also to provide qualified feedback to the developer. Related to the common need for employee qualification is the challenge to hire data talents from the labor market. Company A and C report an intense competition for those data talents but also state to be in a comparably favorable situation due to their strong employer brand.

Fourth, standardization is a key organizational challenge for all companies. While the manufacturing companies A and B require standardized failure codes and KPIs to ensure data integrity and data comparability for analysis, the ICT company C seeks to standardize virtual objects, like customer tickets, to enable an uninterrupted flow across department borders.

Differences

The case study research has also shown four significant differences.

First, the experience in working with data is higher in the ICT company. For Company C, working with data is part of the daily business for many years. Therefore, the company has had much time to built data analytics skills internally. Data utilization, especially in real-time and in big scale, is a trend that has emerged in the manufacturing industry only a few years ago.

Second, Company C has formulated explicitly the objective to use data to inform operational and strategic decision-making and thus to become a *data-driven company.* This clear commitment to data as the foundation for decision-making is unique among the three companies.

Third, while Company A and B have presented use cases concerned with integrating data from different sources, Company C has presented an application that is already exploiting data to improve a key business objective, which is customer satisfaction. In general, the maturity of Company C—in terms of data utilization and especially regarding the use of ML application—is higher compared to the manufacturing companies. The higher maturity of data analytics DBAs in the ICT company is likely to be a result of the first two differences. Years of experience and expertise building in data utilization and a strong commitment to exploit data for decision-making, have enabled Company C not only to develop but already to apply highly advanced DBAs in core activities such as customer support.

Fourth, increased transparency due to increased data collection is perceived mainly as a threat to user acceptance in companies A and B. Transparency on individual performance has a negative touch. As a result, companies A and B have guidelines to restrict the use of operational data to monitor employees' performance. Although Company C also recognizes the threat of employees' rejection of individual performance monitoring, the company at the same time considers the possibility of data-based individual performance assessment as a chance for fair and objective leadership.

6 Consolidation of Findings

This chapter consolidates the results of the quantitative survey of chapter 3, the literature-based findings of chapter 4, and the insights from the qualitative studies of chapter 5. In addition, following the recommendation of Eisenhardt (1989, p. 533), literature was consulted to provide support or contradictions to key findings of the qualitative studies.

This chapter comprises three subchapters. Chapter 0 consolidates the findings regarding DBAs in manufacturing in general. Chapter 6.2 then considers the findings from an LM perspective and discusses the positive and negative implications of DBAs for LM. Finally, chapter 6.3 intends to abstract the findings from the specific use cases and formulates two theoretical implications.

This chapter combines the findings of chapters 3 to 5 to answer the three SRQs on a sound basis. Table 29 provides references to navigate directly to the respective chapter of each SRQ.

Table 29: Sub-research-questions and references to research results

Sub-research-question		**Reference Chapter / Page**
SRQ 1	Which data-based applications exist in manufacturing and what are their objectives?	6.1.2 / 191
SRQ 2	What are key enablers to apply data-based applications?	6.1.3 / 194
SRQ 3	How can data-based applications support lean practices?	6.2.1.2 / 208

6.1 Data-based Application in Manufacturing

This chapter consolidates the findings on DBA in manufacturing from a general point of view and comprises three subchapters. First, chapter 6.1.1 presents five general observations regarding DBAs in manufacturing. Second, chapter 6.1.2 presents the DBAs and their objectives identified in chapter 4.1 and third, chapter 0 presents key challenges and enablers to apply DBAs..

6.1.1 General Observations

This chapter summarizes five general observations of DBAs from literature and praxis.

First, the motivation of manufacturing companies to engage in data utilization is manifold (see chapter 5.4.2.1). Companies can be motivated by the outlook to achieve a particular use case-specific objective, such as increased maintenance effectiveness. However, often the decision to invest in data utilization does not emerge from a concrete need but is driven by the expectation of external stakeholders. Accordingly, top management is pressured to "do something in regard to Smart Manufacturing" and formulate a data utilization strategy. Further drivers are production machinery manufacturers, which equip new machines with sensors as a prerequisite to offer additional services or new pay-per-use models. In doing so, suppliers push the technical prerequisite for DBAs into manufacturing companies.

Second, there is a wide variety of DBAs available fitting the need of the manufacturing industry. As discussed in chapter 4.1, six DBA categories comprising 14 individual DBAs have been derived from a systematic literature review. Those DBA cover a wide range of manufacturing operations, including Planning and Scheduling (I); Production Control (II); Maintenance (III); Internal Logistics (IV); Product Quality Management (V); and Environment, Health, and Safety (VI). Thereby the complexity of DBAs ranges from rather low (e.g., Track and Trace) to very high (e.g., Prescriptive Maintenance). An overview of the six categories and the14 DBAs, including their status of the application in the case companies, is given in the next chapter

Third, by abstracting from the individual objectives of the DBAs, four DBA core functions have been derived in chapter 4.1.5. The four DBA core functions Monitoring (1), Deviation Control (2), Decision Support (3) for humans, and Autonomous Optimization (4) are able to describe the key functionalities of all 14 DBAs. The core function Monitoring is thereby defined as basic core function, as the other three core functions depend on the availability of data collected by this core function. The characterization of the four core functions revealed different levels of value add. While Monitoring is a prerequisite for other core functions, it provides little value as a standalone function. The second core function Deviation Control adds value by controlling and maintaining the status quo. However, from a theoretical point of view, higher value is generated by the core functions Decision Support and Autonomous

Optimization as they not only seek to maintain but to improve the status quo. As the core functions, Decision Support and Autonomous Optimization include advanced data analytics, they are more complex and more challenging to implement than the core functions Monitoring and Deviation Control. Thus, a tendency that higher added value of DBAs comes at the cost of higher complexity can be concluded.

Fourth, the two major data utilization use cases of the manufacturing companies are concerned with facilitating easy and quick access to distributed data. The MAS system of Company A was explicitly driven by the motivation to reduce time and effort to find and access relevant data for diverse data analytics projects (p.125). Before, many of the planned data analytics projects have been cancelled as the effort to collect the data was just too high. The Global Performance Cockpit of Company B is motivated by the outlook to have cross plant data available in near-real time to avoid a negative impact of production stops in one plant on other plants (p.136). Apparently, integrating the diverse data sources into one central data management system is still a major task for manufacturing companies. The ICT company has shown to be more mature in terms of data utilization. While the manufacturing companies still focus on establishing a central database, Company C has already developed and rolled out advanced DBAs to support key objectives. For instance, customer satisfaction, which is a key objective of Company C, is increased by the ML-based Smart Support system.

Fifth, investing in the technical infrastructure for DBAs is getting economically attractive for many companies for the first time. According to the World Manufacturing Forum 2018 (Taisch et al., 2018, p. 15), relevant components such as sensors but also costs for computing and storage capacity have reduced dramatically in the last decade. Budget restrictions were considered a main barrier for digitalization in the 2017 survey presented in chapter 3. Due to the substantial cost reductions, however, it can be assumed that this perception will change in the near future.

6.1.2 Overview and Objectives

This chapter combines the DBA identified in the literature review in chapter 4.1 and the DBAs described in the qualitative studies in chapter 5. Table 30 presents an overview of DBAs identified in the literature along with their objectives and their current state of application in the case companies. Table 30 thus answers the first SRQ.

SRQ 1: Which data-based applications exist in manufacturing and what are their objectives?

The systematic search process of DBAs in the literature is explained in chapter 4.1.2. The overview shows that the range of DBAs covers a broad spectrum of manufacturing activities. Two aspects may indicate the relevance of a DBA in scientific literature and in the manufacturing industry. First, the number how often a DBA was mentioned in a selection of 12 publications (see Table 14) is indicated in the left column of Table 30. Second, the current status in the two manufacturing case companies is indicated in the right column of Table 30.

Table 30: DBAs in theory and practice – overview and objectives

DBA [Number of references]	**DBA Objective**	**Status in Case Companies**
I. Planning and Scheduling		
Production Scheduling [10/12]	Determination of an optimal production plan (e.g., in terms of maximized asset utilization or minimized cycle time) while meeting the constraints in terms of available material and equipment capacity. (p. 67)	Testing phase in companies A and B.
Layout Planning [4/12]	Determination of an optimal production layout (e.g., in terms of low WIP inventory, minimal material handling, and flow orientation) while meeting the constraints in terms of available space and number and type of machines. (p. 69)	Not in use in companies A and B
II. Production Control		
Real-time Control [8/12]	Real-time monitoring of the production process by permanently comparing the actual behavior of the production system to the expected behavior. Real-time notification in case of deviations outside the tolerance. (p. 70)	In use in companies A and B.
System Performance Measurement [2/12]	Transparent overview of the overall system performance by automatic calculation and visualization of KPIs, such as the OEE. KPIs are used to identify trends and for benchmarking against similar machines. (p. 71)	In use in companies A and B.
III. Maintenance		
Condition Monitoring [9/12]	Increase maintenance effectiveness and efficiency by monitoring the equipment health status of machines in real-time. Maintenance activities are triggered only in case of unusual machine behavior. (p. 73)	In use in companies A and B.
Predictive Maintenance [10/12]	Increase maintenance effectiveness and efficiency by monitoring equipment health and forecasting future machine degradation. By forecasting the expected time of machine failure, a maintenance plan is derived that ensures equipment availability while avoiding unnecessary maintenance. (p. 74)	In use in companies A and B
Prescriptive Maintenance [0/12]	Increase maintenance effectiveness and efficiency. The same approach as Predictive Maintenance but in addition, this DBA also derives or even initiates maintenance activities autonomously. (p. 75)	Company A: testing phase Company B: not in use

DBA [Number of references]	**DBA Objective**	**Status in Case Companies**
IV. Internal Logistics		
Track and Trace [5/12]	Traceability of material, container, and products (e.g., position, cycle times, components) by using unique identifiers such as RFID tags. Trace and trace data allow to evaluate a product's history in case of quality issues and increases the transparency of the material flow, thus reducing the level of safety stock. (p. 76)	In use in companies A and B
Material Flow Management [4/12]	Efficient control of the material flow based on real-time demand. Material flow can be controlled either by a production control system, which plans the demand and distribution of material centrally or by a digital pull system. A digital Kanban system detects automatically a need for replenishment and triggers the process by sending virtual Kanban cards. Optimization of the physical distribution of material by data-based path planning for AGVs or milk runs. (p. 77)	In use in companies A and B
Inventory Management [3/12]	Intelligent inventory management to ensure material availability with minimal inventory by increasing the transparency of the current inventory and demands. The objective is to “replace inventory with perfect information.” (p. 78)	In the testing phase in companies A and B
V. Product Quality Management		
Product Quality Monitoring [9/12]	Automatic identification and sorting out of nonconforming products by comparing quality data (e.g., geometrical dimensions) in real-time against reference values. (p. 79)	In use in companies A and B
Product Quality Improvement [9/12]	Systematic identification and preventive avoidance of errors by using collected data on quality issues to perform a systematic root cause analysis. (p. 81)	In use in companies A and B
VI. Environment, Health, and Safety		
Energy Monitoring [7/12]	Reduction of energy consumption, by measuring the consumption of several components and identifying sources of unnecessary energy consumption. (p. 82)	No information on the status quo as DBA was added to the collection after case interviews
Environmental Monitoring [2/12]	Ensure healthy working conditions for employees by monitoring the environmental conditions such as air quality in real-time and compare current condition against reference values. (p. 83)	

6.1.3 Key Challenges and Enablers

This chapter summarizes the key challenges and key enablers found in the qualitative studies. In the first part of this chapter, the challenges are discussed since they are important to understand the enablers presented in the second part.

6.1.3.1 Key Challenges

6.1.3.1.1.1 Overview

Table 31 provides an overview of the key challenges of using DBAs.

Table 31: Challenges of data utilization in manufacturing (emerged in qualitative studies)

Category	Main Challenge	Sub-challenge	Page
Employees (E)	C-E.1. Initial Resistance of Employees	C-E.1.1 Fear of Job Loss C-E.1.2 Fear of Loss of Job Autonomy C-E.1.3 Fear of Control	168
	C-E.2. Employee Qualification	C-E.2.1 Shop floor Employees C-E.2.2 Middle Management C-E.2.3 Top Management	169
Organization (O)	C-O.1. ROI Calculation of Investment Decisions	C-O.1.1 Uncertainty of Results C-O.1.2 Chicken and Egg Problem	170
	C-O.2. Misuse of Data	C-O.2.1 Internally C-O.2.2 Externally	171
	C-O.3. Comparability of Data	C-O.3.1 Within the Plant C-O.3.2 Within the Manufacturing Network	171
	C-O.4. Access to Expert Knowledge	C-O.4.1 Limited Internal Resources C-O.4.2 Competition for Data Scientists	172
Technology (T)	C-T.1 Basic Requirements	C-T.1.1 IT System Performance C-T.1.2 Data Security C-T.1.3 Technical Infrastructure[1]	172
	C-T.2 Distributed Data	C-T.2.1 Data Integration C-T.2.2 Data Access	173
	C-T.3 Miscellaneous	C-T.3.1 Insufficient Data C-T.3.2 Inapt Algorithms	174
Key	C-E.1.1: sub-challenge 1 of main challenge 1 of the category **E**mployees [1]Identified as key requirement of DBAs based on use cases from literature (see chapter 4.1.6)		

This chapter consolidates key challenges discussed in the case studies and the expert interviews with the key DBA requirements identified in chapter 4.1.6. As shown in Table 31, a variety of challenges for data utilization in the context of manufacturing exists.

Some of the challenges discussed are DBA context-sensitive while others have been indicated by several partners. DBA context-specific challenges are, for instance, the need to translate implicit human knowledge to explicit machine-readable knowledge (Company C), and the challenge of insufficient data to evaluate the effectiveness of a predictive maintenance DBAs.

Context independent challenges have emerged in all three main categories. Overcoming initial resistance of employees was mentioned as consistently as a key challenge as ensuring the right level of employee qualification. Similarly, a performant IT system and data security were identified unison as challenges for data utilization. Finally, the impossibility to seriously forecast the results of data analytics DBA and, thus, the organizational challenge to deal with the unpredictable ROI was a reoccurring theme in the discussions.

Table 31 lists eight main challenges plus two additional and rather specific sub-challenges assigned to the challenge C-T.3 Miscellaneous. When also considering the key requirements of DBAs identified form the literature in chapter 4.1.6, three more challenges need to be added: technical infrastructure (1), data availability (2), and know-how (3).

Taking a closer look, however, reveals significant overlaps of the three additional challenges to the challenges listed in Table 31, hence allowing the integration of the key requirements into the key challenges. The requirement technical infrastructure can be logically assigned to the main challenge C-T.1 Basic Requirements. The requirement data availability is covered by the main challenge C-T.2 Distributed Data, and the requirement know-how can be merged into the main challenge C-E.2. Employee Qualification.

The remaining eight main challenges are briefly summarized below. More details can be found by following the reference to the respective page in chapter 5.4 in the right column of Table 31. By following the references, each challenge can be traced back to the respective section in the case study or expert interview.

6.1.3.1.2 EMPLOYEES

6.1.3.1.2.1 Initial Resistance of Employees

Ensuring employees' acceptance of new DBAs and overcome initial resistance is the first key challenge. Employee resistance may be provoked for three reasons. First, fear of job loss due to the perception that DBAs are more efficient than humans and will gradually replace human labor (C-E.1.1). Second, fear of loss of job autonomy.

The case studies have shown that employees fear to lose the freedom to derive self-determined decisions, if they have to work with a decision support system, such as the *Smart Support* (Company C) system or a predictive maintenance application (C-E.1.2). Third, fear of control is triggered due to the constant expansion of the data collection. Manufacturing data can be misused to control the behavior of individuals on the shop floor (C-E.1.3).

6.1.3.1.2.2 Employee Qualification

Employee qualification is the second key challenge that was identified consistently by all interview partners. The first challenge is to define the required level of competence of each employee and the second challenge is to provide target-group-specific training. During the qualitative studies the critical competencies of three different groups of employees were outlined: shop floor employees (C-E.2.1), middle management (C-E.2.2) and top management (C-E.2.3).

The consolidation of key requirements of DBAs (see chapter 4.1.6) comprises primarily technological requirements but also *know-how*. According to Table 17, the implementation and use of DBAs require multidisciplinary skills, including IT know-how, manufacturing domain expertise as well as expertise that is specific to the DBA. For instance, predictive maintenance requires specific skills such as data analytics skills to detect patterns and forecast future equipment health. In conclusion, literature and qualitative research indicate the challenge of building the broad set of skills, which are necessary to implement and use DBAs.

6.1.3.1.3 ORGANIZATION

6.1.3.1.3.1 ROI Calculation of Investment Decisions

From a managerial perspective, a key challenge is the infeasibility to forecast the ROI of investment decisions in data analytics DBAs. While traditionally investment decisions are based on an ROI calculation, this concept is hardly applicable for data analytics DBA projects. The underlying problem is that the result of those DBAs is not predictable before the analysis (C-O.1.1). The situation is characterized as a *chicken and egg* problem. Managers wait for a proof of concept before granting project funds, but the proof of concept cannot be provided prior to the analysis. Overcoming the chicken and egg problem is a key organizational challenge *(C-O.1.2)*.

6.1.3.1.3.2 Misuse of Data

Data protection and avoidance of data misuse is a second organizational challenge. Misuse of data can occur internally and externally. The case companies and the researchers have pointed out that tracking more and more data on the shop floor allows relatively precise control, not only of employees' performance but also of their behavior. As discussed above, *fear of control* is a key driver of employee resistance against DBAs. Missing personal data protection can even cause delays in or blockade of projects that collect and process personal data (*C-O.2.1*).

Protection of data from external misuse is also a central theme for the case companies. This is in line, with the requirement ***data protection*** derived in chapter 4.1.6. Whereas data security is mostly considered as a technological challenge, the discussion with practitioners revealed that it is often not the IT system but employees which are the weakest part of data protection (*C-O.2.2*). Thus, avoiding data theft is not only a technological but also an organizational challenge.

6.1.3.1.3.3 Comparability of Data

Many DBA not only process data of one machine but of several machines in one factory (*C-O.3.1*) or even from different factories in the production network. A key organizational challenge is to ensure the comparability of data from those different sources. For instance, the Global Performance Cockpit of Company B provides a near real-time overview of KPIs from different sites. To allow meaningful comparisons, the KPIs' definition and calculation need to be consistent (*C-O.3.2*).

6.1.3.1.3.4 Access to Expert Knowledge

According to the case companies, the internal resources of data analytic expertise are limited. For manufacturing companies, the field of data analytics is rather new, but even the ICT company struggles to build enough internal data analytics expertise as expressed by the quote: “we know that we need more data-affine employees” (C-O.4.1). This finding is in line with the survey results presented in chapter 3. To the question, what are barriers to use digital, the survey participants reported, that shortage of manpower and employee qualification are both among the top three barriers (see Figure 12). Although digitalization and data utilization are not equivalent, the quantitative study results support the conclusion, that building expert knowledge for new applications internally remains a key organizational challenge.

An alternative to building the required skills internally is to hire data experts from the job market. According to a study by Macuvele et al. (2018, p. 39), almost one out of two companies is currently hiring or intends to hire data experts. The high demand for data experts and the scarce availability on the labor market results in a “war for data talents.” Finding and hiring sufficient external data experts, despite lucrative offers from other industries, is a challenge not only for the manufacturing companies but also for the ICT company (C-O.4.2).

6.1.3.1.4 TECHNOLOGY

6.1.3.1.4.1 Basic Requirements

Based on the literate-based use cases of DBAs discussed in chapter 4.1.4, key requirements have been summarized in Table 17 in chapter 4.1.6. Most of these requirements are technological requirements and are therefore consolidated with the technology challenges identified in the qualitative studies in chapter 5.

The first basic technological requirement is to create the basis for using data in a specific DBA. Therefore, a technical infrastructure is required to collect, transfer, preprocess and protect data. Data collection poses the challenge to collect data in high quality, thus demanding for automated data collection with accurate and reliable sensors. Real-time availability of data is critical for DBAs such as *Real-time Control*. Due to the lack of a high-performing mobile data network (e.g., 5G standard), ensuring a robust data transfer with very low latency times is currently a key technological challenge. Data preprocessing is identified as another challenge due to the fact that data in manufacturing systems is often unstructured and originates from several sources. Finally, data protection is critical to avoid data theft of criminals. A survey among 126 companies has shown that data security concerns hampers the introduction of digital technologies in SMEs (see chapter 4.1.6.1).

In contrast to the literature, the qualitative studies focused more on employee and organizational challenges than on technological challenges. Nevertheless, IT system performance and data security are described as fundamental requirements. IT system performance is necessary to transfer and process data in reasonable time. High computing power is necessary for some DBAs; for instance, when including complex simulations (C-T.1.1). Regarding data security, the interview partner did not go into technical details but emphasized that data security is not only a technological challenge. Employees need to be briefed to avoid data theft by passing on information unintentionally (C-T.1.2).

6.1.3.1.4.2 Distributed Data

Finding and merging distributed data has caused enormous effort in Company A and thus made many data analytics uneconomical. This situation motivated the introduction of the *Manufacturing Analytics Solution* that provides a single platform for data access. Two challenges are related to the implementation of such a system. First, to integrate data from various sources (C-T.2.1) and second, to provide an interface that allows authorized employees easy and user-friendly access to the required data (C-T.2.2).

The challenges of the category *Miscellaneous* (C-T.3) are use case-specific and are therefore not considered as a general key challenge.

6.1.3.2 Key Enablers

6.1.3.2.1 OVERVIEW

Table 32 presents an overview of the enablers of DBAs and links the enabler to a related challenge. Thereby, this chapter answers the second SRQ.

SRQ 2: What are key enablers to apply data-based applications?

Table 32: Enablers of data utilization in manufacturing

Enablers of Data Utilization in Manufacturing

C	Main Enabler	Sub Enabler	Related Main Challenge	Related Sub-challenge	P
Employees	E.1. Foster Acceptance	E.1.1 Understand the Value of DBAs	C-E.1. Initial Resistance of Employees	C-E.1.1	175
		E.1.2 Being Part of the Solution Development		C-E.1.2	176
	E.2. Role-specific Training	E.2.1 Basic Training of End-users	C-E.2. Employee Qualification	C-E.2.1	176
		E.2.2 Citizen Data Scientists			177
		E.2.3 Advanced and Holistic Understanding at the Middle Management Level		C-E.2.2	178
		E.2.4 Basic but Holistic Understanding at the Top Management Level		C-E.2.3	178
Organization	O.1. Create Favorable Conditions	O.1.1 Structured Approach to Select DBAs	C-O.1. ROI Calculation of Investment Decisions	C-O.1.1	179
		O.1.2 Ensure Management Buy-in		C-O.1.1 & C-O.1.2	180
		O.1.3 Put the Right Leader in Place		C-O.1.1	180
		O.1.4 Maintain the Motivation		C-O.1.1	181

Enablers of Data Utilization in Manufacturing					
Organization	O.2. Data Usage Transparency	O.2.1 Data Guidelines	C-O.2. Misuse of Data	C-O.2.1 & C-E.1.3	182
		O.2.2 Data Security Officer		C-O.2.1 & C-O.2.2	182
	O.3. Standardization	O.3.1 Management Responsibility for Standardization	C-O.3. Comparability of Data	C-O.3.1 & C-O.3.2	182
	O.4. Internalize External Expertise	O.4.1 Convince with non-monetary Benefits	C-O.4. Access to Expert Knowledge	C-O.4.2	183
		O.4.2 Benefit from Cooperation		C-O.4.1	183
Technology	T.1. State-of-the-art Technology	T.1.1 Benefit from Falling Component Costs	C-T.1 Basic Requirements	C-T.1.1 - C-T.1.3	183
		T.1.2 Cloud Computing			184
Key:		E.1.1: Sub enabler 1 of main enabler 1 of the category **E**mployees, C: Category P: page reference to detailed enabler description in chapter 5.4.2.3.			

This chapter consolidates the enablers discussed in the qualitative studies and, if applicable, links the enabler to one of the key challenges presented in the previous chapter. Thereby, readers can learn how companies have addressed those challenges in a real-world context. In doing so, managers can benefit from the documentation of the challenges and enablers and apply the gained insights for similar problems in their companies (Gassmann, 1999, p. 11).

As depicted in Table 32, seven main enablers and 17 sub enablers have been derived in the qualitative studies. Although a majority of the key challenges are addressed by corresponding key enablers, no complete coverage was reached. Due to the lack of interview partners with a background as technician or in IT, the technological challenges are less addressed than the challenges of the categories Employees and Organization.

The seven key enablers are presented briefly in the remainder of this chapter. Like in the previous subchapters of this consolidation chapter, links to the detailed enabler description are provided on the one hand in the overview Table 32

Table 32 and on the other hand at the end of each enabler brief description. From there, further references are given to navigate to the text passage in the case studies or expert interviews in which the enabler was introduced and discussed originally.

6.1.3.2.2 EMPLOYEES

According to a study conducted by the Fraunhofer IAO, human labor in industrial production will continue to be of central importance in the coming years (Spath et al., 2013, p. 45). With regard to the interaction between DBAs and human employees in production, the two challenges *Resistance of Employees* and *Employee Qualification* have been identified. The corresponding enablers to address these challenges are *Foster Acceptance* and *Role-specific Training*.

6.1.3.2.2.1 Foster Acceptance

The case studies have revealed three risks for employee's acceptance of DBAs: *Fear of Job Loss* (C-E.1.1), *Fear of Loss of Job Autonomy* (C-E.1.2), and *Fear of Control* (*C-E.1.3*). The corresponding enabler *Foster Acceptance* comprises two components. First, employees need to understand the value of DBAs and second, employees need to be involved in developing a new solution.

Understand the Value of DBAs

To support employees' understanding of a DBA's value, the management needs to communicate clearly how the application is beneficial for the company (e.g., increasing the competitiveness by decreasing costs and thus securing jobs). Thereby the *Fear of Loss of Job* is reduced. In terms of data quality, understanding the implication of entering flawed data into the system for the outcomes, increases the motivation of end-users to invest extra effort into entering data correctly (E.1.1).

Being Part of the Solution Development

Involving employees in the development of a new solution has been a critical enabler for the DBA use cases in the case studies. For instance, maintenance employees refused to work with a predictive maintenance DBA that was developed without their involvement. Service employees showed resistance against a data-based decision support tool, as they felt they are losing the autonomy to develop the best solution to a customer's problem on their own. Involving the maintenance employees in the development of the predictive maintenance application and granting the service employees the ability to overrule the decision support system has shown to be effective to reduce the *Fear of Loss of Job Autonomy* (E.1.2).

The third cause of employee resistance, *Fear of Control,* is discussed as part of the enabler *Data Usage Transparency* in subchapter 6.1.3.2.3 (organizational enablers).

A comparison of the results on employee resistance with Figure 13 in chapter 3 shows a strong consensus. Figure 13 visualizes the barriers of using digital technologies to support lean. However, as data utilization is closely linked to the emergence of digital technologies, a comparison between the quantitative (survey) and qualitative findings is considered worthwhile.

Interestingly, employee resistance was one of the main barriers to digitalization for SP companies, while it was only a minor barrier for the overall sample. This observation implicates a limitation for the generalizability of the finding that employee resistance is a key challenge. All case study companies are large and technological mature companies, the perspective of SME companies, however, is not covered in the case studies. It might be, that challenges like budget and infrastructure restrictions (see Figure 12) are a more significant challenge for companies in the initial phase of digitalization. Only if these barriers are overcome and concrete applications are implemented, like in the case companies, the challenge of employee resistance gains importance.

Hirsch-Kreinsen et al. (2018, p. 176) have reviewed several studies on the future of work in a digitalized production environment and found conflicting results. While some studies are optimistic and expect a general upgrading of industry jobs, other studies refer to the risk of job loss and the risk of reduced competencies. Furthermore, they conclude that the potential for employee control is a central risk of digital technologies (Hirsch-Kreinsen et al., 2018, p. 180). These findings are in line with the three risks for employees' acceptance of DBAs presented above.

6.1.3.2.2.2 Role-specific Training

All partners of the qualitative studies have confirmed that the right employee qualification is one, if not the key challenge to implementing and using DBAs effectively. Although chapter 4.1.6 lists primarily technical requirements, it does comprise know-how as essential requirement for DBAs. The qualitative studies not only identified employee qualification as key challenge but also highlighted the fact that different roles in the organization require different skills. Thereby, the following three groups are differentiated DBA end-users (1), who are usually working on the shop floor; managers from the middle management that are responsible for implementing DBAs (2); and top managers (3), who often initiate and sponsor DBA projects.

The corresponding enabler to this challenge is *Role-specific Training.* To be effective, training needs to be tailored to the roles and requirements of employees. This finding is backed by the German state secretary of the Federal Ministry of Education and Research, who argues that job-specific qualification is decisive for industry 4.0 (Kagermann, Wahlster, & Helbig, 2013, p. 56).

Basic Training of End-users

End-users require basic training, which conveys the skill to operate the new application. However, end-users are likely to use a new application more effectively, if they do not only know how to use the application but also understand the basic logic beyond the tool. Furthermore, a basic understanding of the underlying logic enables end-users to participate in the development process of new DBAs by providing qualified feedback to DBA developers. Providing basic training that allows end-users involvement in the development of a new application contributes to a better adaption to the actual user needs and higher acceptance among end-users. (E.2.1)

Citizen Data Scientists

Several authors highlight the crucial relevance of manufacturing domain expertise; for example, a detailed understanding of a process or a machine for using manufacturing data effectively (Åkerman et al., 2018, p. 416; Lee et al., 2014, p. 7; Liao & Wang, 2013, p. 229; Mayr et al., 2018, p. 625). Manufacturing domain expertise is, for instance, necessary to formulate accurate data models as well as interpreting data analytics results. Therefore, employees that combine years of experience on the shop-floor with data analytics skills are vital for the effective exploitation of manufacturing data. To enable employees to build and combine both skills, the concept of developing *Citizen Data Scientists* has emerged as promising approach.

Citizen Data Scientist are technicians or engineers that already work within the company in the manufacturing area and are selected for extra training in data analytics skills. *Citizen Data Scientist* candidates usually have a strong intrinsic interest and motivation in data but also have exiting IT skills, often from a personal interest. After intensive training—involving formal training, self-training, and exchange with data and IT experts—*Citizen Data Scientists* serve as internal consultants for DBA projects on the shop floor (E.2.2).

Advanced and Holistic Understanding at the Middle Management Level

Although DBA projects are often initiated by the top management, the responsibility for actually implementing a DBA is usually passed on to the middle management. Ensuring adequate qualifications of middle managers is especially challenging, due to the variety of skills required to manage DBA projects, including manufacturing system and process know-how, change management skills, and an advanced understanding of data analytics.

There is no one single enabler that ensures the full range of required qualifications. However, a two-pillar training system for middle managers was presented in the case studies. The first pillar comprises formal training in a dedicated 150-hours data analytics course. The second pillar fosters internal knowledge exchange. Selected employees can work part-time in an IT or data analytics team for up to 40 weeks, thus learning new skills on the job. (E.2.3)

Basic but Holistic Understanding at the Top Management Level

A lack of technological expertise on higher management levels has been identified as a challenge for companies intending to apply DBAs (C-E.2.3). Due to the lack of time, it is unlikely that top managers will attend a full content formal course as the one discussed above. However, top managers can benefit from a course with reduced scope and depth, tailored to their requirements. By understanding the basic concepts of data analytics, top managers are better qualified to evaluate DBA proposals, either from within the organization or from external partners (E.2.4).

6.1.3.2.3 ORGANIZATION

6.1.3.2.3.1 Create Favorable Conditions

The qualitative studies have revealed four organizational enablers that increase the probability of successful DBA projects.

Structured Approach to Select DBAs

First, a structured approach facilitates the identification of the most promising DBA opportunities and ensures adequately resource endowment. The structured approach, called *Analytic Business Case Review* in the case study, includes intense discussions between the DBA expert and the process owner to evaluate the potential of DBAs for the specific context. The extensive experience of the process owner of the process is essential for the evaluation (O.1.1).

Hence, the *Analytic Business Case Review* is in line with the propositions of the *CRISP-DM* framework presented in chapter 2.5, which proposes that understanding the context of the intended analysis is the first step of the iterative CRISP-DM cycle.

Ensure Management Buy-in

Second, management buy-in essential at the start of a DBA project, to grand project resources even if the ROI on these investments is highly uncertain. However, it is also critical after the initial phase, especially if the DBA project takes longer as planned or does not deliver the intended results in the first iteration.

Companies that are persistent even in case of setbacks tend to be more successful in implementing DBA projects (O.1.2).

Put the Right Leader in Place

Third, the ideal leader of DBA projects fulfills the following requirements. He or she has sufficient time to manage the project, has relevant expertise in manufacturing, IT and data analytics, is persistence in case of setbacks, and has a certain level of seniority to ensure access to the management, access to resources, and the ability to break initial internal resistance (O.1.3).

Maintain the Motivation

Fourth, to keep the motivation of all stakeholders, including project members and management high, easy to achieve *Quick Wins* should be designed and build-in in DBA projects. Those little success stories along the way “keep people happy and management calm,” even if the main objective has not yet been achieved (O.1.4).

6.1.3.2.3.2 Data Usage Transparency

Data Guidelines

The risk misuse of personal data is a general barrier to DBAs. To address employees' fear of control, companies are well-advised to create full transparency over the collection and use of data within their plants. Discussions with employee representatives support companies to formulate data guidelines which are transparent and protect individual data (O.2.1).

Data Security Officer

In addition, companies may appoint a Data Security Officer who is responsible for ensuring compliance with external and internal data and privacy policies (O.2.2).

6.1.3.2.3.3 Standardization

Management Responsibility for Standardization

To use failure codes to identify patterns, these failure codes need to be standardized. The same applies to KPIs within a plant and across plants. Aligning historically differently defined metrics and codes is resource-consuming and sites show little intrinsic motivation to do so. As a result, it takes active leadership that enforces standards within and across production sites (O.3).

6.1.3.2.3.4 Internalize External Expertise

Convince with non-monetary Benefits

Companies struggle to meet the internal demand for employees with data analytic skills. Besides the internal qualification of employees, two other options are available: to hire data experts externally or to cooperate with external partners. To hire data professionals despite the intense competition of solvent financial companies and consultancies, manufacturing companies need to attract candidates with non-monetary benefits. One key argument to convince potential candidates is the perspective to test and evaluate ideas and prototypes in the real world, close to the working space, and with instant feedback (O.4.1).

Benefit from Cooperation

Sourcing in external expertise can alternatively be achieved by collaborating with external partners (e.g., consultancies and research institutions). Universities offer access to human resources and expertise and are usually less expensive than consultancies. (O.4.2)

6.1.3.2.4 TECHNOLOGY

An enabler to meet the challenges *IT System Performance* and *Technical Infrastructure*, is to invest in state-of-the-art technology.

6.1.3.2.4.1 State-of-the-art Technology

As discussed in chapter 6.1.3.1, an essential requirement for DBAs is to install the technical infrastructure required to collect, transfer, pre-process, and protect data.

Benefit from Falling Component Costs

According to Figure 12 of chapter 3, *budget restrictions* are an important barrier to use digital technologies. Assuming that the same applies to DBAs, providing sufficient financial resources to install the required technical infrastructure is a barrier, at least for smaller companies. However, this challenge is mitigated by an external enabler, which is the fact that components such as sensors, computing power, and storage capacity are getting less expensive rapidly (T.1.1).

Cloud Computing

Part of the enabler *State-of-the-art Technology* is cloud computing. As shown in chapter 2.3.3, cloud computing is a core technology of SM and allows to use and pay exactly the amount of computing and storage capacity needed. Access to scalable resources is cost-effective and more flexible than traditional fixed IT resources (T.1.2). In addition to cost and flexibility aspects, Vogel-Heuser et al. (2017, p. 135) highlight the high robustness of cloud computing. Data remain accessible, even in the case than one server is temporarily down as the data is stored at multiple servers at the same time.

6.2 Implications for Lean Manufacturing

6.2.1 Support of LM by Data Utilization

The conjunction of SM and LM manufacturing was discussed in chapter 2.4. Three different perspectives are currently debated in the literature. First, LM as the foundation for SM. The key argument for this perspective is that SM needs robust, transparent, and standardized processes as a foundation. Second, SM advances LM (e.g., by increasing the flexibility of the LPS). The third perspective argues that SM and LM are mutually beneficial without concluding the direction of support. For instance, scholars have found empirical evidence that SM technology maturity and LM maturity are positively correlated.

The work of this dissertation is in line with the second perspective. However, instead of researching the impact of SM or its related concepts of industry 4.0 and digitalization on LM as a whole, this work has focused on the impact of data utilization, as part of SM, on LM. The evaluation was done on two different levels. The case studies have shown that industry representatives tend to think about and discuss the impact of data utilization on lean more in general terms instead of actual applications. Thus, the following chapter 6.2.1.1, summarizes the expected main benefits on a higher level.

This dissertation, however, has set the target to go beyond general implications of data utilization on lean, but to conduct a systematic evaluation of the impact of actual DBAs on established lean practices. Thus, the research follows the recommendation of Mayr et al. (2018, p. 623), who have criticized that most papers addressing the impact of SM technology on lean stay on a general level, not linking the impact to a particular lean method. The results of the DBA – lean practice evaluation is summarized in chapter 6.2.1.2.

6.2.1.1 Main Benefits: Full Transparency and Decision Support

As part of the case study interviews, the case company representatives were asked to evaluate the impact of data utilization on LM. Without dedicated questions on the impact of DBAs und lean practices, the answers remained on a general level.

The two manufacturing companies agree that LM will not change fundamentally due to the opportunities of data utilization. The five basic lean principles *define value from the customer perspective*, *identify the value stream*, *flow*, *pull*, and *strive for perfection* (see chapter 2.2.1.2) will remain guiding principles. However, the tools and methods that enable companies to follow these lean principles may be impacted due to data utilization.

The case studies had a strong focus on company-specific use cases of data utilization. The use cases were selected by the interview partners based on their perception of its relevance compared to other use cases. Both use cases of the manufacturing companies seek to provide easy and fast access to distributed data with the key objective of increasing transparency of the on-site production systems as well as enabling production network transparency. Especially concerning LM, increased transparency is the main benefit of data utilization expected. Increased transparency of the current inventory and demand will be used to identify and reduce unnecessarily large inventories. Furthermore, transparency and full access to data will support the lean CI process by supporting CI tools that rely on accurate data.

“Gaining direct data-based insights on where the biggest potentials for optimization and the next improvements are, as opposed to relying on opinions and prioritizing them.” Company C goes beyond the goal of creating transparency to support lean but seeks to gain data-based decision support for improvements. Data-based decision support is also the underlying concept of the Smart Support use case. This use case perfectly illustrates the potential of DBAs to support key lean objectives, namely creating customer value and reduce waste. Based on documented legacy customer problems and solutions, the Smart Support system proposes solutions with the highest likelihood of solving the customer problem sustainably in minimal time. Thus, the system increases customer value due to fast and qualified customer support. Furthermore, by defining the sustainability of the solutions as a decisive factor, waste due to recurring problems is eliminated.

6.2.1.2 Support of Lean Practices by DBAs

This subchapter summarizes the findings of chapter 4.2 *Impact of Data-based Applications on Lean Practices* and thus answers the third SRQ.

SRQ 3: How can data-based applications support lean practices?

The evaluation followed a systematic approach, comprising four steps.

First, 10 widely established lean practices were identified from the literature. Therefore, a collection of lean practices compiled by Shah and Ward (2003) as result of a systematic literature review was taken as a basis. To reduce the number of lean practices, only the nine most cited ones were selected from a total of 21 lean practices. This selection was complemented by the lean practice Value Stream Mapping, which emerged after 2003 but is today considered to be a major lean practice (see chapter 2.2.3).

In a second step, 14 DBAs were identified in a systematic literature research (see chapter 4.1.2). The 14 DBAs are assigned to six DBA categories, as visualized in

Figure 15. In the third step, the concept of a DBA – Lean Practice Impact Matrix was introduced in chapter 4.2.2.

The actual DBA – lean practice evaluation was performed in the final fourth step. To avoid a bias of the results due to personal and subjective perceptions and assumptions, the evaluation procedure also consists of four steps, including literature research, theoretical reasoning, academic feedback, and feedback from industry professionals (see

Figure 16).

The result of the pairwise evaluation is already consolidated in an overview presented in the Lean Impact Matrix in Table 19 in chapter 4.2.2.2. A detailed discussion of the impact evolution is given in the following chapter 0. The remainder of this chapter summarizes the six top DBA - lean practice combinations with expected high or at least moderate support potential.

6.2.1.2.1 LEAN PRACTICE PREVENTIVE MAINTENANCE

The lean practice Preventive Maintenance aims to minimize unexpected breakdowns by performing maintenance preventively. Fewer machine breakdowns lead to higher process stability. Traditionally, Preventive Maintenance usually follows a periodic maintenance plan with fixed intervals.

Predictive Maintenance is an important application of data utilization in manufacturing. It is not only discussed frequently in academic articles but is also seen by practitioners as a promising approach to increase equipment availability and process stability. The application *Predictive Maintenance* uses current real-time equipment data as well as historic machine defect data to predict the remaining useful life. Thus, *Predictive Maintenance* mitigates the risk of too much or too little maintenance of traditional preventive maintenance approaches, which follow a periodic, fixed-interval maintenance approach. By focusing on necessary maintenance activities, not only equipment availability is increased, but also the costs for unnecessary maintenance activities are reduced. In consequence, *Predictive Maintenance* can considerably support the lean practice Preventive Maintenance to be performed more effectively (fewer machine breakdowns) and more efficiently (reduction of maintenance effort).

All interviewed companies report having *Predictive Maintenance,* at least to some extent, already in use. A barrier for a broader implementation are high costs for *Predictive Maintenance,* which in many cases still exceed the costs for changing spare parts routinely by following a fixed plan. In general, however, there is little doubt that *Predictive Maintenance* will be increasingly relevant and will eventually advance to the new standard of maintenance.

6.2.1.2.2 *LEAN PRACTICE QUALITY MANAGEMENT*

The lean practice Quality Management is a part of the broader lean practice TQM. The objective of the lean practice Quality Management is to minimize product failures by identifying faulty products and apply failure root-cause analysis for systematic failure reduction.

The DBA *Quality Monitoring* identified defective parts or products by comparing quality parameters with reference values in real-time. One example of data-based *Quality Monitoring* is visual inspection with a camera system combined with an image processing algorithm. In case of deviations, the respective part is sorted out automatically. Hence, the application *Quality Monitoring* minimizes the risk of defective parts being passed on to the process step or even to the customer and thus support the lean practice Quality Management

However, an even stronger positive impact on the lean practice Quality Management is expected from the DBA *Product Quality Improvement*. Literature suggests that by monitoring and collecting quality data, systematic root cause analyses can be conducted faster and more systematically and thus more effectively. As a result, factors for product quality can be identified and optimized. To this end, ML is a promising technique to classify and detect failures but also to identify their defect root causes.

6.2.1.2.3 *LEAN PRACTICE CONTINUOUS FLOW PRODUCTION*

The lean practice Continuous Flow Production aims to create a situation in which the production components flow through the value creation process without interruptions and waiting times.

Continuous Flow Production can be supported from two directions. First, maintenance DBAs minimize the likelihood of unexpected machine breakdowns and therefore impact process stability positively. Process stability, in turn, is essential for Continuous Flow Production.

Second, although process stability is crucial for Continuous Flow Production, the positive impact of a higher process stability is restricted if other requirements, such as timely material supply are not met. To increase the reliability of the material supply, the DBA *Track and Trace* is used to identify the exact location of material or container in real-time. Real-time data on material and current material position enable the DBA *Material Flow Management* to optimize the material supply (e.g., by timely replenishment of required parts or by smart AGV path planning).

The combination of equipment reliability and material supply reliability is strongly supporting the lean practice Continuous Flow Production.

6.2.1.2.4 *LEAN PRACTICE PULL/KANBAN*

The lean practice Pull/Kanban promotes a demand-oriented material flow. In contrast to a push system, the production of new parts is triggered only be the need for replenishment of the next downstream process step. Following the lean practice Pull/Kanban, therefore, limits the amount of WIP inventory in the process. Pull is most often realized by a Kanban system. Traditionally, Kanban works with physical Kanban cards.

Popular in academic literature is the concept of a digital Kanban system, also called e-Kanban. In contrast to the traditional Kanban system, e-Kanban can automatically evaluate if replenishment is necessary and uses digital instead of physical Kanban cards. Scholars have identified several advantages of e-Kanban compared to traditional Kanban, including, faster transmission of replenishment signals, no lost cards, and the ability to adjust the lot size dynamically.

The advantages of e-Kanban are confirmed by the manufacturing case companies. Consequently, they already have replaced traditional paper-based Kanban with e-Kanban.

E-Kanban can be assigned to the DBA *Material Flow Management.* Hence, by eradicating weaknesses of traditional Kanban systems, this DBA has a strong potential to support the lean practice Pull/Kanban.

6.2.1.2.5 *LEAN PRACTICE VALUE STREAM MAPPING*

The lean practice Value Stream Mapping documents processing and waiting times, as well as the flow of material and the flow of information to visualize the current status of a production process.

The lean principle may benefit considerably from integrating real-time data from manufacturing into the value stream map. Literature and practitioners acknowledge the potential of real-time data-enhanced *Value Stream Mapping,* also referred to as VSM4.0, to draw a more precise picture of the actual value stream. Especially regarding the trend for increased product variety and customization, one interview partner argues that the value streams of similar but slightly different products may not be perfectly identical. Manufacturing data can reflect these minor differences as well as dynamic changes in the value stream—for instance, different bottleneck situations depending on the product. Furthermore, having the necessary data available allows for the conducting of VSM regularly with low human effort.

6.2.1.2.6 *LEAN PRACTICE CONTINUOUS IMPROVEMENT*

The lean practice Continuous Improvement focuses on the improvement of processes, services, and products. CI activities often follow a rigorous scientific approach and include the use of systematic tools such as the DMAIC cycle. The DMAIC cycle includes an analyze phase which benefits from access to accurate manufacturing data. Due to the availability of several kinds of manufacturing data, such tools can not only be applied more effectively but also faster and with less effort. Hence, data availability supports the lean practice Continuous Improvement.

The relevant data are collected by multiple DBAs including *Real-time Control*, *Condition Monitoring*, *Track and Trace*, *Product Quality Monitoring*, and *Energy Monitoring.* Before a problem can be addressed by CI methods, however, it needs to be detected. By providing automatically calculated KPIs and visualizing their trend development, the DBA *System Performance Measurement* facilitates the identification of hidden problems within the production system. Furthermore, full transparency of performance indicators allows internal and external benchmarking with comparably low effort due to automatic data collection and KPI calculation.

In summary, all DBAs collecting data support the lean practice Continuous Improvement, but the strongest support potential has the DBA *System Performance Measurement.* This conclusion is consistent with the view of the industry representatives, who argue that transparency will be a key enabler for LM in general and for CI in particular.

6.2.2 Potential Threats

The industry representatives' expectation of the impact of data utilization on lean is consistently positive. Also, a majority of the DBA – lean practice combinations have either indicated a positive or no impact. Nevertheless, chapter 4 and chapter 5 have also shown at least two potential threats for LM arising from data utilization. Chapter 6.2.2.1 outlines the risk of alienation from the basic concept of lean due to the omnipresent availability of manufacturing data. Chapter 6.2.2.2 sketches a possible incompatibility of the current technology selection process in LM with the challenge of the uncertainty of results of DBA investments.

6.2.2.1 Alienation from the Basic Concept of Lean

The suggestion that the lean philosophy is not compatible with information technology is advocated by lean purists for many years (Sugimori et al., 1977). Among the possible conflicts are push vs. pull and simplicity vs. IT complexity (see chapter 2.2.4).

When taking a critical position to the impact of data utilization on LM, three risks for alienation from the basic concept of lean can be concluded based on the findings of this dissertation.

6.2.2.1.1 Value Stream Mapping 4.0 vs. Go and See

The concept of VSM4.0 has aroused both interest and concern from practitioners. The positive implications have already been discussed in the previous chapter. On the negative side, industry professionals see the risk of substituting the physical presence on the shop floor as part of VSM with remote access to manufacturing data.

Consistently highlighted is the importance of being “where the action happens” to gain a sound understanding of the real process. Accordingly, process data without process understanding is of little value and bears the risk of data misinterpretation. The basic principle of VSM is described as “go and see.” Therefore, a virtual value stream created by remotely accessible data would contradict the fundamentals of VSM. Nevertheless, a balanced integration of a few metrics is desirable as long as it does not replace shop floor presence.

Meissner et al. (2018, pp. 83–84) shares the critical view on automated data collection and KPI calculation and highlights the risk that the “automation of data collection and processing bears the risk that shop floor workers feel disconnected from the performance measurement process and therefore lead to alienation from shop floor management.”

6.2.2.1.2 HIGH COMPLEXITY VS. EMPLOYEE INVOLVEMENT AND CI

As discussed in chapter 2.2.3.2.4, ensuring employee involvement in CI is a key task of the management in an LPS.

A threat to employee involvement in CI arises from the high complexity of some DBAs. For instance, the Smart Support system is based on ML. In the data science community, there is currently a discussion, if modern ML systems are *black boxes* or only very complex and thus merely impossible to understand (Card, 2017). Either way, it becomes almost impossible for employees to fully understand how an ML system derives its results. However, without understanding the cause and effects in a system, employees may be discouraged of or even deterred from thinking about how the system can be improved. CI will not be anybody's responsibility anymore, as advocated by lean, but the job of a group of highly specialized ML experts. Kieviet (2016, p. 41) summarizes this threat by asking: “How do you identify waste if the chaos is optimized digitally?”

6.2.2.1.3 INFORMATION OVERFLOW VS. SIMPLICITY

Lean promotes simplicity (Maguire, 2016, p. 32), which means focusing on the essentials. Eiji Toyoda, a former CEO of Toyota, already pointed out in 1983 the risk of information overflow to the ability to solve problems:

> *“Society has reached the point where one can push a button and be immediately deluged with technical and managerial information. This is all very convenient, of course, but if one is not careful there is a danger of losing the ability to think. We must remember that in the end it is the individual human being who must solve the problems.” Eiji Toyoda, 1983, cited in Bell (2006, p. 0)*

The risk of information overflow is today higher than at any time before history, due to the omnipresent collection and accessibility of information. Meissner et al. (2018, pp. 83–84) found that one disadvantage of digital technologies for shop floor management is the risk of tracking too many KPIs due to the high availability of data and the automatic KPI calculation. Hence, following the lean philosophy, managers need to resist the temptation to track too many KPIs, as a high number of KPIs maks it difficult for employees to focus on the essential information.

6.2.2.2 Technology Selection Without Proven Added Value

The state of research chapter on LM has included a discussion on the role of technology in LM (see chapter 2.2.4). The key message of Liker (2004) regarding the integration of new technology in the TPS is that every new technology has to prove its ability to support established value-creating processes before it is bought and installed.

As suitable as this approach is for technologies like automatization technology, it is not fully applicable for technologies relying on data analytics, e.g., ML-based decision support systems. As shown in chapter 5.4.2.2, a distinct characteristic of data analytics DBAs is the uncertainty of its results. Although the same application has worked successfully in a different context, here is no guarantee that it will deliver meaningful results in another context and with other data.

As a consequence, the motto *value demonstration first, investments later,* may implicate a very hesitant position of LPS managers in terms of taking the risk of investing in data analytics applications. On one hand, this position reduces the risk of non-rentable investments, but on the other hand, chances for significant improvements will be missed.

The potential of data utilization in manufacturing projected by scholars (Kusiak, 2017; O'Donovan et al., 2015b, p. 1; Tao et al., 2018; Wuest et al., 2016), the variety of fields of application for DBAs (see Figure 15), and the real-world examples of added value by DBAs in the case studies, however, lead to the conclusion that a strict rejection of *risky* investments in DBAs will not support a manufacturer's competitiveness in the long run.

6.3 Theoretical Implications

6.3.1 The Investors Dilemma of DBAs

The three case companies, which are all large companies and leaders in their respective businesses, have invested significant resources in building the foundation for data utilization and are already using several DBAs.

However, taking a look at Figure 13 in chapter 3 indicates that these SP companies are not representative of the manufacturing industry in general. Quite on the contrary, Figure 13 shows that almost every second company in the study has equipped less than 20 percent of their equipment with sensors for real-time monitoring. Only one out of eight companies claim to have more than 80 percent of the equipment monitored. Furthermore, collecting data is not equivalent to using data. To this end, a 2017 survey of the ITEM-HSG and the RWTH Aachen among 100 companies has revealed that only a small share of 5.5 percent of the available data, which corresponds to 11 percent of the collected data, is actually used (Wenking, Benninghaus, & Groggert, 2017, p. 35) (see Figure 18).

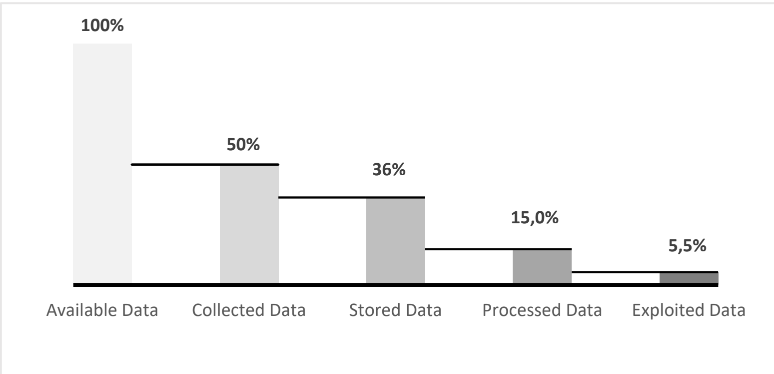

Figure 18: Share of exploited data of available data (based on Wenking et. al [2017, p.35])

Given the fact that only 11 percent of the collected is exploited in some way, one can say that data utilization in the manufacturing industry is not a success story so far. In appears that apart from major companies, the manufacturing industry, in general, is hesitant to invest in data exploitation. This observation motivated to investigate the causes of this investment reluctance.

A potential reason was indicated in the case studies and expert interviews. The outcomes of data analytics DBAs projects are highly afflicted with uncertainty and a serious ROI forecast of these DBA projects is almost impossible. The uncertainty has two dimensions. First, prior to analyzing a set of data, it cannot be guaranteed that meaningful patterns will be found and second, if patterns are found, the resulting added value of exploiting the pattern is difficult to determine.

The problem of uncertain outcomes, which also leads to the chicken and egg problem, is not equally important for all DBAs. Only data analytics DBAs that go beyond the first two core functions of data monitoring and deviation control (see chapter 4.1.5) but include actual data *analysis* are afflicted. This applies in particular to the DBAs *Predictive* and *Prescriptive Maintenance* (e.g., to identify wear patterns), *Material Flow Management* (e.g., to optimize AGV path planning), and *Product Quality Improvement* (e.g., to identify critical to quality factors).

To better understand the challenges of investments in data utilization, the following chapter 6.3.1.1 evaluates the ratio of needed effort and expected return in terms of the added value of a selected DBA. Afterward, chapter 6.3.1.2 builds on the preliminary observations and adds the aspect of the uncertainty of results into consideration. The chapter presents the phenomena of the *Valley of Tears* in DBA investments, which might serve as a theory-based explanation for the reluctance of investments in data utilization discussed above.

6.3.1.1 Three Levels of Added Value of DBAs

Inspired by the observation of Shao et al. (2014, p. 2194) that *influence* has a higher value than *observe*, or in other words, *decision support* is more valuable than *data collection*, this chapter evaluates the added value of a DBA in three different stages of its lifecycle. In addition to the value, the required effort is assessed based on the requirements and challenges of DBAs identified in chapter 4.1.6 and chapter 5.4.2.2 respectively. As a representative of all data analytics DBAs, the DBA *Predictive Maintenance* was selected for evaluation.

The result of the evaluation is visualized in Figure 19. Accordingly, each stage has a distinct level of added value for the company.

Level 1: The first level is labeled *Transparency* as the value proposition of the DBA on this stage is to *make things visible.* For the exemplary DBA, the objective of the first stage is to monitor the equipment condition to create transparency about the current equipment health status. In this first stage, the DBA comprises the first core function of data monitoring.

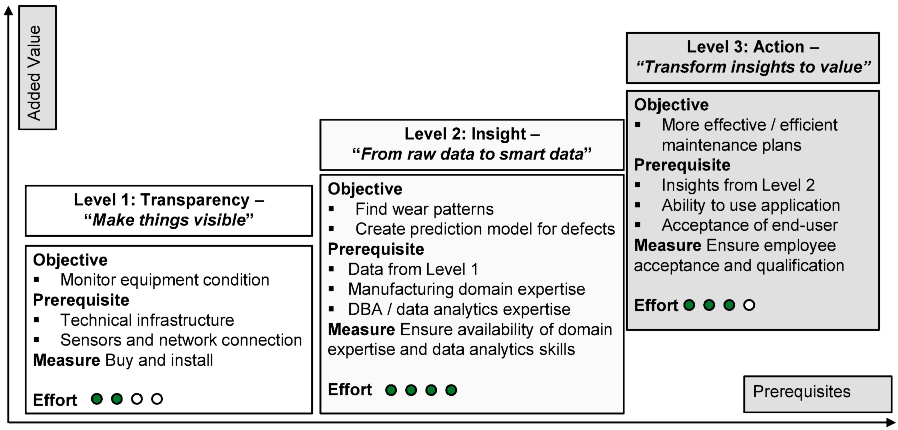

Figure 19: DBA value model

Value: By creating transparency over the equipment health status, the DBA creates some but limited value to the company on the first stage. However, only by having the equipment condition data monitored, the condition is not improved, hence the added value at this stage is low.

Effort: The effort of installing the DBA at this basic level is estimated to be comparatively low. Condition monitoring requires a technical infrastructure that includes at least sensors and network connections. However, the infrastructure can be acquired externally and the installation can be supported by the supplier, thus reducing the internal effort for providing the necessary prerequisite of level 1.

Level 2: The second level is labeled *Insight* as the value proposition is to transfer *raw data to smart data.* At this stage, data analytics is applied. Referring back to the example, the objective of the second stage is to find patterns in machine data and, if patterns are found, create a prediction model for machine defects.

Value: During the analysis, as long as no patterns are found, the added value compared to level 1 is near zero. And even if patterns are found in the data, the actual added value in terms of achieving the DBAs objectives of higher machine uptime is not existent yet. Hence, the added value of level 2 compared to level 1 is marginal.

Effort: To perform data analytics the DBA on level 2 requires the data output of level 1 as a prerequisite. Furthermore, on this stage, a particularly high degree of expertise know-how is needed. On one hand, data analytics expertise is necessary to perform the analysis, manufacturing domain experts are needed to interpret the data, and additional specialists (e.g., for ML) may be required to create and refine the ML algorithm of the DBA. As the combination of these skills and expertise is rare and experts are often costly, companies need to invest significant effort in the DBA on this level. Also, unlike sensors, not all the expertise can be bought externally, but needs to be built internally over a long period.

Level 3: The real value creation starts not before level 3, which is labeled *Action.* Accordingly, the value proposition of level 3 is to *transform insights to value.* The objective of the exemplary DBA is to exploit the patterns from the second level to derive more effective and efficient maintenance plans.

Value: Only on the third level, the DBA provides significant additional value by actually exploiting the data and found patterns to serve a certain purpose that is beneficial for the company. In the exemplary case, the DBA *Predictive Maintenance* is now able to support the derivation of more effective and efficient maintenance plans, thus reducing the number of machine breakdowns, increasing machine productivity, and minimizing maintenance costs.

Effort: As seen in the qualitative studies in chapter 5, to use a new application effectively, companies need to ensure end-users' acceptance of the application and invest in additional employee qualification. That is the reason why the required effort on stage 3 is still high. It is though below the effort required on level 2, as the experts are not needed to use the tool.

Discussion of Assumptions with Senior Researchers

The *DBA value model* rests on several assumptions and has not been tested with real-world use cases for generalizability. However, key assumptions of the model have been discussed with both senior researchers. Based on their personal experience, they confirm the plausibility of the following underlying assumptions of the model.

Assumption 1: Significant value for the company is only created if the findings of data analysis are exploited for a certain purpose. Hence, the added value of level 3 is substantially higher than the added value on level 1 and level 2.

Assumption 2: The required effort investment reaches its maximum on the second level. The main driver for the high effort on this level is the required expertise, which cannot be acquired externally in the short term (Professor Schuster).

Assumption 3: Effort exceeds added value on the second level by far.

Assumption 4: Effort for exploiting insights (level 3) is lower than performing the actual data analysis (level 2). This assumption is confirmed to be plausible for manufacturing applications. The fourth assumption may not apply, however, regarding novel data-based business models (e.g., *not ownership* models), as they introduce new complexity while companies currently have little experience with these business models.

6.3.1.2 The Valley of Tears in DBA Investment Decisions

This chapter builds on the assumptions and implications of the DBA value model and adds the identified challenge of high uncertainty of data analytics outcomes into consideration.

For deriving an investment decision, the course of value and effort over the DBA lifetime is decisive. Mastering level 1 and level 2 is a precondition for the third level. Hence, added value and required effort have to be considered cumulatively over all three stages of a DBA's lifecycle.

Figure 20 depicts two graphs. On the left side, the implications of the DBA value model are visualized. As the expected value of the DBA on the first value level is low, it is exceeded by the effort needed to install the physical infrastructure. However, both the added value and the required investments are comparably low. Hence, on the first value level, a small loss is generated per time unit. c, effort increases significantly, while the added value compared to level 1 is marginal. Consequently, entering the second level leads to a significantly higher loss per time unit compared to level 1. The ratio between effort and value, however, is inversed on the third level. When the insights of level two are transformed into value-creating actions, considerable value is created for the first time. At the same time, effort is reduced compared to value 2. Thus, the model predicts a strong surplus of value created against effort invested on the third level.

The surplus is represented by the green area labeled *Profit* in the left part of Figure 20.

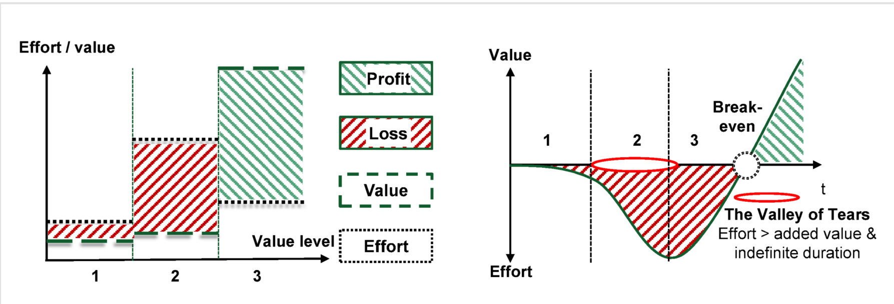

Figure 20: The Valley of Tears in DBA investment decisions (own illustration)

The right part of Figure 20 visualizes the course of the cumulated value and effort over time. In the initial phase, the DBA is on the first value level. According to the left part of Figure 20, the effort slightly exceeds the value in this phase. During the time the DBA remains on the first level, the consolidated surplus of effort is growing slowly. A DBA project enters the critical phase after the infrastructure is installed on the first level and the actual data analysis starts on the second level. Due to the highly negative value to effort ratio, every time unit spent on the second stage causes a significant loss. This loss keeps on cumulating until the third level is reached.

Situations in which effort is higher than the return are common to all companies. What is critical, however, is the fact that the *uncertainty of results*, makes it hardly possible to estimate the duration of the second stage seriously. Hence, companies face the dilemma of losing money at every time unit on the second stage, while the duration of this stage is indeterminable.

Prof. Wuest instinctively called this situation the *Valley of Tears*, which serves in the following as a metaphor to describe the unfavorable situation on the second stage.

The remainder of this chapter contrasts traditional investment calculation against the investment calculation of data analytic DBA projects. Both situations are depicted in Figure 21

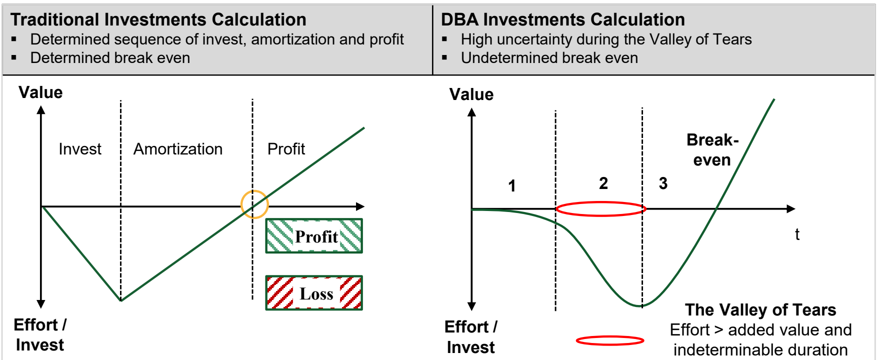

Figure 21.

Figure 21: Traditional investment vs. DBA investment calculation
(own illustration, left part based on Fleig (2019))

The left side describes traditional investment calculations, comprising the three phases of investing, amortization, and profit. The key difference between both situations is that traditional investment decision-making relies on fairly accurately determinable time periods until an initial invest starts to generate value. Thus, the beginning of the amortization period is predictable. Furthermore, the value of the investment, for instance, in terms of reduced costs, can be estimated. Hence, the expected amortization time and thus the break-even point can be projected.

In comparison, DBA investment decisions pose higher requirements to the management in their investment decision-making process. The time until the third value level is reached is underdetermined and thus the break-even point is undetermined too.

The *Valley of Tears* poses three significant risks to DBA projects.

First and directly derivable from

Figure 21, the DBA investment calculation is highly afflicted with uncertainty. The break-even point cannot be estimated seriously. Hence there is the risk of remaining in the second value level, which would implicate a constant high surplus of effort over created value. Typically, managers tend to prefer to invest resources in projects with a reliable ROI. Therefore, DBA projects may have a structural disadvantage in competing for project funds.

Second, data analytics DBA may not deliver the expected results within the expected time. Risk-averse managers may react to initial setbacks by canceling the project. However, the qualitative studies have also shown that data analytics is an iterative process and results may be found only in the second or third iteration. Hence, the conditions of the *Valley of Tears* can foster the cancelation of projects which have already made large progress to stage 3, leading to sunken costs, no results, and demotivated project members. To avoid early termination, Professor Wuest suggests building in *Quick Wins* along the way (see Enabler O.1.4).

Third, the uncertainty of the *Valley of Tears* can also cause the opposite effect. Companies may keep on going to spend money on a project which permanently fails to deliver the expected results, maybe just for the simple reason that there are not usable insights hidden in the data.

The conscious or unconscious decision to avoid the three risks by abandoning DBA projects is a plausible explanatory approach for the low degree of data utilization in manufacturing.

6.3.2 Technology Acceptance Model: Evaluation and Propositions for Extension

The TAM model is a user acceptance theory that intends to explain the adoption of new technologies. In its original form as introduced by Davis (1989) the TAM postulates that the *perceived usefulness (U)* and the *perceived ease of use (EOU)* are the key determining factors of users' technology acceptance (see Figure 3). Thereby, a technology has a high *perceived usefulness (U)* if the user feels that the technology enhances his or her job performance and a high *perceived ease of use (EOU)* if the user believes that using the technology is free of effort.

The number of case studies of this dissertation is too small to confirm the propositions of the TAM empirically. However, one use case has been discussed that indicates a missing element of the TAM.

Table 33 contrasts the expected acceptance according to the TAM and the actual end-user acceptance for this DBA use case, in two different situations.

Table 33: Technology Acceptance Model evaluation

No.	Situation	U	EOU	Acceptance according to TAM	Actual end-user acceptance	PIC
1	Decision Support system (initial situation)	+	+	+	Initial employee rejection	-
2	Decision Support system (current situation)	+	+	+	Employee acceptance and technology usage	+

Key: (U) perceived usefulness, (EOU) perceived ease of use, (PIC) perceived individual contribution

The first situation described is inspired by the Smart Support system use case. The *perceived usefulness (U)* is rated as positive as the employee is supported to give fast und sustainable feedback to customer requests. Also, the *perceived ease of use (EOU)* is positive, as the new system is described as more user-friendly than the earlier versions. According to the TAM, two positive evaluations of these key factors should result in a high user acceptance of the system.

However, the initial situation was characterized by a rejection of the system by some users, as they saw the system as a threat to their job autonomy. The fear was triggered by the perceived risk of being downgraded from a self-determined solution developer to a decision system order receiver (see 5.2.4.4). The obvious contradiction between the prognosis on user acceptance and the actual user acceptance causes the assumption that the TAM is missing a critical element.

The very right column of Table 33, labeled (PIC) stands for *perceived individual contribution* and is defined by the authors of this dissertation as “the degree to which a person believes that his or her contribution is relevant for success.” *Perceived individual contribution* is proposed as an additional factor to explain user acceptance of new technology and is motivated by the reaction of the case company to initial employee resistance to its new decision support system.

The company distinctly communicated that the end-user will “remain in the driver’s seat.” This means that service employees always have the opportunity to overrule the system and communicate a solution to the customer, which is different from the one suggests by the system. Although a majority of the system's suggestion is accepted by end-users, having the final decision power, significantly increases the *perceived individual contribution* to the final solution. Today, under these circumstances, the decision support tool is broadly accepted and welcomed as support to enhance the individual job performance. This situation is represented in situation 2 in Table 33.

The positive impact of integrating end-user in the solution development process on the user's acceptance of the solution has been identified as an enabler to overcome employee resistance (see Enabler E.1.2). The argument that *perceived individual contribution* is important for employees' motivation to use a tool or system is supported by the famous work: “A theory of human motivation” by Maslow (1943). Accordingly, all people strive for achievements, recognition, and appreciation (Maslow, 1943, p. 381), whether in a private or business context.

Based on the observations in this case study, the author proposes to extend the basic TAM model by the factor *perceived individual contribution (PIC),* as shown in Figure 22.

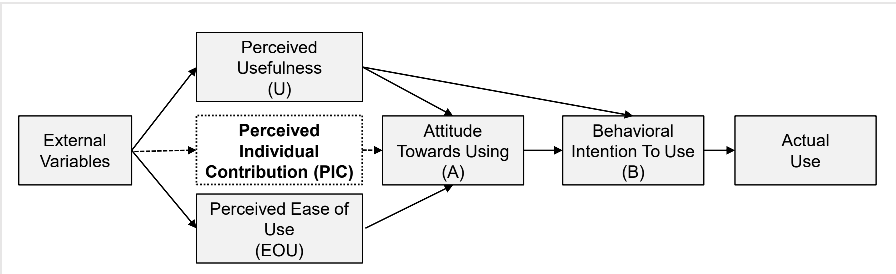

Figure 22: Proposition for extended Technology Acceptance Model (adapted from Davis, 1989)

7 Conclusion and Outlook

This chapter provides a summary of the key results of this dissertation. Chapter 7.1 briefly summarizes the key results and answers the MRQ. Chapter 7.2.1 consolidates the theoretical contribution, followed by the practical contribution in chapter 0. Chapter 7.3.1 points out the limitations of this research, while chapter 7.3.2 gives an outlook on potential further research.

7.1 Summary of Results

LM is described as the currently most influential manufacturing paradigm. However, besides LM the newly emerging concept of SM, also referred to as industry 4.0, receives increased attention from academia, media, and government. Hence the question arises how these two production paradigms fit together (see chapter 1.1). Closely linked to the emergence of SM is the increasing attention for data utilization in manufacturing (see chapter 2.3). A 2017 study on the future of lean, found that companies perceive big data as a digitalization trend with a strong potential to support LM. Furthermore, the survey results illustrate than mature lean companies are investing more effort into real-time data monitoring and, on average, perform data analytics on a higher maturity level (see chapter 3.2).

A subsequent literature review identified a lack of integrated consideration of LM and data utilization as a means to enhance LM. To address this gap, the following MRQ was formulated: “How can manufacturing companies be enabled to implement DBAs to support lean practices?”

To answer the MRQ, three SRQs are derived. While the first and second SRQ focus on the manufacturing industry in general, the third SRQ bridges the existing gap between LM and data utilization. The first SRQ aims to create an overview of existing DBAs and their objectives in manufacturing. The second SRQ investigates key enablers for implementing DBAs in a real-world context. At last, SRQ 3 links the previous research to LM by evaluating how DBAs can support established lean practices. As the SRQs have already been answered in detail in the consolidating chapter 6, only a short summary of the results of each SRQ, including references is given below.

SRQ 1: Which data-based applications exist in manufacturing and what are their objectives?

The first SRQ was answered based on a systematic literature search (see chapter 4.1.2). In total, six DBA categories, including 14 DBAs have been identified and visualized in Figure 15. The six DBA categories are Planning and Scheduling (I); Production Control (II); Maintenance (III); Internal Logistics (IV); Product Quality Management (V); and Environment, Health, and Safety (VI). Table 30 in chapter 6.1.2 provides an overview of the individual DBAs including their objectives and their current status in the two manufacturing case companies. The overview shows a broad variety of DBA objectives, addressing several distinct functions in manufacturing.

SRQ 2: What are key enablers to apply data-based applications?

The second SRQ was answered based on the results of qualitative studies. The qualitative studies include three case studies, among which are two describing leading manufacturing companies and one portraying the situation of a leading ICT company. To increase the validity of the results, the case studies were complemented by two expert interviews with senior researchers in the field of data utilization in manufacturing. To answer SRQ 2, the discussions with industry and academic partners addressed not only enablers but also challenges of applying DBAs.

Based on DBA use cases identified in the literature, key requirements were summarized in chapter 4.1.6. Most of these requirements are technological requirements addressing the need for suitable technical infrastructure and data availability. The qualitative studies, however, have demonstrated that it would be unwise to consider data utilization only from a technological perspective.

The cases and expert interviews have revealed a variety of critical challenges and enablers within the categories Employees and Organization. In total, this dissertation has identified and discussed nine main challenges, comprising 21 sub-challenges and seven main enablers, comprising 17 sub enablers. The challenges are consolidated in chapter 6.1.3.1, and an overview is given in Table 31. Key challenges related to employees are *Initial Resistance of Employees* and ensuring role-specific *Employee Qualification.* Organizational challenges are the uncertainty of the *ROI Calculation of Investment Decisions*, the risk of *Misuse of Data*, *Comparability of Data* and *Access to Expert Knowledge*. From a technology perspective, challenges are to meet the technical *Basic Requirements,* such as the technical infrastructure, and to integrate *Distributed Data.*

As far as applicable, enablers identified in the qualitative studies are linked to identified challenges. Thus, readers can learn how companies have dealt with challenges in a real-world context. The enablers are consolidated in chapter 6.1.3.2, and an overview is given in Table 32, also referencing the addressed challenges.

Key enablers related to employees are to *Foster Acceptance* and provide *Role-specific Training*. Organizational enablers are to *Create Favorable Conditions* for data utilization projects, to ensure *Data Usage Transparency*, the *Standardization* of metrics KPIs and codes and to *Internalize External Expertise.* Enablers from a technology perspective are *Falling Component Costs* for the technical infrastructure and the access to scalable computing and storage resources via *Cloud Computing.*

By emphasizing the importance of employee and organizational challenges and enablers, the case companies show, that the industry has learned its lessons from the CIM era in the 1980s and early 1990s. Brandt (2017) has summarized implications from the CIM era for industry 4.0 and quotes Braun, Förster, and Vorspel-Rüter (1988, p. 21): "Practical experience shows that IT problems are not the key challenges for bringing the CIM philosophy to reality. First and foremost, personal and organizational challenges have to be overcome." [31]

SRQ 3: How can data-based applications support lean practices?

The current discussions in academia on the impact of SM on LM tend to remain on a general level and lack to address the impact on actual lean tools and methods (see chapter 1.4.2). To address this weakness in current research, the third SRQ evaluated the impact of DBAs on widely applied lean practices. The underlying assumption of this research indicated in the research framework (see Figure 2) is that by supporting lean practices, DBAs indirectly enable LM to achieve its objectives (e.g., waste reduction) more effectively.

Ten lean practices were selected (see chapter 2.2.3) and considered for potential impacts by the 14 DBAs identified in chapter 4.1. The methodology followed a pairwise evaluation approach, resulting in the DBA – Lean Practice Matrix, showing all identified impacts in a concise overview (see chapter 4.2.2). The evaluation followed a four-step procedure (see Figure 16). Accordingly, the following six lean practices are likely to benefit substantially from the support of DBAs.

First, the lean practice Preventive Maintenance will benefit particularly from the DBA *Predictive Maintenance,* due to more effective and efficient maintenance plans. Second, the lean practice Quality Management is likely to be supported by the DBAs *Quality Monitoring*, by detecting and sorting out flawed products, and especially by the DBA *Product Quality Improvement*, by exploiting legacy data on quality problems to perform systematic root cause analyses. Third, the lean practice Continuous Flow Production can profit from a high machine and process stability, due to effective maintenance thanks to the maintenance DBAs. Also, a stable and efficient material

[31] Translated from German

flow, ensured by the DBA *Material Flow Management* has a positive impact on Continuous Flow Production.

Fourth, existing deficiencies of the traditional Pull/Kanban system are mitigated by e-Kanban and smart material supply route planning, both features of the DBA *Material Flow Management.* Fifth, the lean practice Value Stream Mapping can be performed faster, more accurate, and with less human effort by integrating accurate and near-real time manufacturing data. Especially, dynamic changes and small differences in the value stream due to a high product variety can be reflected better by integrating automatically collected data. This approach is referred to as VSM 4.0 in the literature.

Supporting DBAs are all DBAs with the core function of data monitoring, namely *Real-time Control*, *Condition Monitoring*, *Track and Trace*, *Product Quality Monitoring*, and *Energy Monitoring*.

Sixth, very similar to Value Stream Mapping, the lean practice *Continuous Improvement* is very likely to benefit from the easy to access availability of accurate and (near) real-time data. For instance, several CI tools rely on data for different types of analysis (e.g., the DMAIC circle). Whereas all monitoring DBAs listed above are contributing to data availability, the application *System Performance Measurement* is seen as most promising DBA to support CI. By providing full transparency of the performance of the manufacturing system by automatically collected metrics and calculated KPIs, negative trends and areas for improvement are detected faster and more systematically. While a potential support of DBAs is expected for most of the lean practices of the categories TPM, TQM, and JIT (see chapter 2.2.3), no potentials to support the human-centered lean practices Cross-functional Work Force and Self-directed Work Teams were found.

Although the DBA – lean practice combinations with a positive support potential dominate, the evaluation has also found potential negative impacts. Also, the qualitative studies and reviewed literature have brought to light potential threats of data utilization in manufacturing, which are especially critical for LM companies. The threats are consolidated in chapter 6.2.2. The first potential threat is *Alienation from the Basic Concept of Lean* due to the temptation to rely on computer provided data instead of actually understanding the process behind the data by going to where the actions happen and see the process firsthand. The second threat arises from the uncertain value generated by DBAs, which is incompatible with the traditional lean approach to select technologies.

By consolidating the findings of the three SRQ, MRQ is answered.

How can manufacturing companies be enabled to implement data-based applications to support lean practices?

SRQ 1 has identified 14 DBAs in six categories, which cover a wide range of functions in manufacturing. Then a systematic evaluation as part of answering SRQ 3 has shown at least six opportunities to enhance the effectiveness of widely used lean practices by applying DBAs. Hence, companies can build on this systematic evaluation when striving to bring their LPS to a higher level. Investments in data utilization infrastructure, software, and people might be easier to justify by referring to positive impacts of data utilization on the established LPS. However, as shown by answering SRQ3, implementing DBAs successfully poses several challenges to companies. Hence, manufacturing companies can profit from the collection of enablers identified in a real-world industry context to overcome these challenges. Understanding that data utilization is not only a technological challenge but needs significant effort to overcome employee-related, and organizational challenges is critical to implement and use DBAs. Instead of repeating the failure of the CIM era to focus only on technological aspects, alleged soft factors such as employee acceptance and qualification as well as data protection have to be emphasized from the start.

7.2 Contributions

Contributions of this research are twofold, impacting both academicians and practitioners. Chapter 7.2.1 consolidates the theoretical contribution of this dissertation, followed by a summary of the practical contributions in chapter 0.

7.2.1 Theoretical Contributions

This dissertation has four main theoretical contributions.

First, it has consolidated literature from different streams, including Big data, DM, ML, and SM to compile a structured collection of DBAs for manufacturing (see chapter 4.1.3). Furthermore, based on the identified DBAs, a systematic classification system, assigning fourteen DBAs to six distinct DBA categories was developed and presented (see Figure 15).

Second, there is a plethora of literature focusing on LM or in different ways on data utilization in manufacturing; however, there is a lack of academic literature reporting the effects of the implementation of both. This lack was addressed by evaluating how data utilization can support LM. Thereby, this dissertation follows the request to perform the evaluation on a tangible level, which is to evaluate the impact on actual lean practices. To structure the analysis and visualize the results, the DBA - Lean Practice Impact Matrix was developed in chapter 4.2.2, following a four-step pairwise evaluation process of DBA – lean practice combinations.

Third, based on two separate studies conducted by the ITEM-HSG in 2017, the conclusion was drawn that except large and mature organizations, manufacturing companies tend to be hesitant to invest courageously into data utilization. Building on different insights from the qualitative studies, this dissertation proposes an explanatory approach that may contribute to understanding the hesitant position of many manufacturing companies regarding investments into data utilization. Based on the *Valley of Tears* observation, three risks inherent to data analytics projects are described, leading to the so-called *Investors Dilemma of DBAs* (see chapter 6.3.1).

Fourth, observations from the case studies were used for a critical evaluation of the TAM. Based on a contrary example, the explanatory power of the TAM was questioned and a suggestion for extension was made. In addition to the essential influencing factors for technology acceptance proposed by the TAM model, *perceived ease of use* and *perceived usefulness*, this dissertation suggests also integrating the *perceived individual contribution* of a user to the solution of a certain problem into the model (see chapter 6.3.2).

7.2.2 Practical Contributions

From a practical perspective, this dissertation makes three main contributions.

First, the compilation of DBAs from the literature may not only be interesting for academics but also for the industry. Production managers can benefit from the collection by getting a feeling about the wide range of improvement opportunities provided by DBAs. By contrasting the objectives of the DBAs with the company's current most pressing needs, the identification of suitable DBAs is facilitated. References to use cases in the literature allow industry managers to deep dive if further details are required.

Second, based on the qualitative studies, nine main challenges, comprising 21 sub-challenges of data utilization, have been derived from a real-world context. Studying a problem in its natural setting is a key argument for conducting case study research as the findings tend to have a higher practical relevance. Managers value case study research as they like to learn from the experience of other companies, especially regarding arising problems and strategies to overcome this problem. Hence, documentation of the challenges is very valuable for these kinds of managers.

However, even more valuable than understanding the challenges other companies have faced is to understand how the companies have overcome or at least addressed these challenges. Following this argumentation, this dissertation has linked the identified enablers to the respective challenge. As shown in the enabler overview in Table 32, the enablers include tangible recommendations for actions, such as introducing the concept of Citizen Data Scientists and starting the discussion with labor representatives to formulate guidelines on data utilization everybody is aligned with.

The collection of key enablers demonstrates that setting the focus on technological aspects exclusively is the wrong way to make data utilization a success story in a company. Decision-makers who are too young to remember the failure of the CIM era can especially learn from this dissertation that technology is only one among several critical challenges for data utilization, including employee challenges and organizational challenges.

Finally, this dissertation can give lean managers new impulses to enhance the performance of their company's LPS. As discussed in the introduction to this work, companies face the challenge of steadily increasing customer demand for high quality, low prices, and high variant flexibility. At the same time, participants of the 2017 lean study have reported decreasing productivity gains by standard lean tools and methods, leading to the conclusion that the low hanging fruits have already been picked. By documenting support potential of at least six widely applied lean practices, this work can contribute to a stronger perception of data utilization as an enabler and driver of higher LPS performance.

Despite the fact, that the DBA – lean practice impact evaluation has found more potential for positive than negative impacts on lean, this dissertation has also shed some light on potential threats for LM due to data utilization (see chapter 6.2.2). Hence, lean managers are sensitized to cautiously balance the degree of data utilization to benefit from the support potentials without risking the alienation from the basic concept on lean.

7.3 Limitations and Further Research

7.3.1 Limitations

In the context of this research, four limitations have to be pointed out. First, motivated by the intention to reduce and sharpen the scope of this dissertation, only DBAs applicable within the boundaries of a production site were considered for the DBA collection in chapter 4.1. However, the case studies have shown that such an artificial separation does not match real-world conditions. Companies operating more than one production site are eager to use the data across the network, for instance for internal coordination or internal performance benchmarking. Hence, by excluding network-oriented DBAs, this work cannot claim to have encompassed all DBAs relevant to today's manufacturing industry.

Second, regarding the results of the qualitative studies, the reader should bear in mind that only a limited number of three case studies have been conducted. All companies are headquartered in the German-speaking area, thus causing the risk that findings are not equally applicable to other regions. For example, the very strong emphasizes of the protection of individual data is likely to be unique for central Europe, especially compared to the United States and China. To mitigate the limitation of a small number of cases, the industry experience of the senior academics was added as an additional source of information in the qualitative studies.

Third, all three case companies are large organizations and leaders in their respective businesses. In this respect, the question arises if the findings are transferable to SME companies. An indicator that differences exist are findings of the Lean2020 study shown in chapter 3.2. While the SP companies (of which all are large companies) perceive employee resistance as a key barrier to digitalization, other barriers such as budget restrictions are more relevant according to the overall sample, comprising many SME companies.

The fourth limitation arises from the fact that the DBA value model presented in chapter 6.3.1 was not challenged and refined with further DBA use cases. The underlying assumptions of the model have been derived from the findings of this research but are not firmly grounded in the literature. Hence, the plausibility of these assumptions can be questioned. To compensate for the lack of theoretical grounding, the key assumptions were discussed with both senior researchers.

7.3.2 Further Research

From the author's perspective, the following three aspects are worthwhile for further research. First, as stated in the previous section on limitations, the restriction on DBAs operating within the boundaries of the production site excludes several very promising DBAs that rely on data exchange between two or more players. Further research is advised to evaluate the potential of DBAs in one or more of the following areas of application. To improve the performance of the production system from a production network perspective, to optimize the supply chain by fostering the integration of supplier data or to better specify customer value by integrating customer data. For instance, DBAs integrating customer data may reveal insights on the most important product features, thus allowing to deploy R&D funds specifically on these features. Arguably, by opening the scope of DBA, new challenges will emerge. A critical challenge raised in discussions in the course of this work is how to organize data exchange with stakeholders outside the own organization, without disclosing critical information.

The second suggestion for future research is motivated by the identified threats for the basics of lean. As shown by the Smart Support use case ML, in the current public discussion rather called AI, has made the transition from the developing and piloting phase to the integration into the daily business. It is reasonable to assume that AI will play an increasingly important role in almost all areas, including manufacturing, in the near future. Thus the contradiction of lean which is advocating simplicity and understanding of cause and effects, and AI, which will be perceived as a *black box* by many people, is likely to pose a complex challenge to LPS managers. Further research might address the question how companies can profit from the opportunities of AI without risking to lose the ability to understand how AI decisions are derived.

Finally, further investigations are required to better understand the mechanisms underlying the investment decisions for data utilization projects. The explanatory approach suggested in this dissertation, called the *Investors Dilemma of DBAs*, is only a very initial description of challenges for investment decision-making inherent to data analytics DBA. To ensure the validity of the approach, key assumptions of the DBA value model need to be tested in further studies.

References

Agard, B., & Cunha, C. D. (2007). Manufacturing Plant Layout Supported with Data Mining Techniques. In S. Tichkiewitch, M. Tollenaere, & P. Ray (Eds.), *Advances in Integrated Design and Manufacturing in Mechanical Engineering II* (pp. 373–383). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-1-4020-6761-7_25

Ajzen, I., & Fishbein, M. (1973). Attitudinal and normative variables as predictors of specific behavior. *Journal of Personality and Social Psychology*, *27*(1), 41–57. https://doi.org/10.1037/h0034440

Åkerman, M., Lundgren, C., Bärring, M., Folkesson, M., Berggren, V., Stahre, J., . . . Friis, M. (2018). Challenges Building a Data Value Chain to Enable Data-Driven Decisions: A Predictive Maintenance Case in 5G-Enabled Manufacturing. *Procedia Manufacturing*, *17*, 411–418. https://doi.org/10.1016/j.promfg.2018.10.064

Albliwi, S. A., Antony, J., & Lim, S. A. h. (2015). A systematic review of Lean Six Sigma for the manufacturing industry. *Business Process Management Journal*, *21*(3), 665–691. https://doi.org/10.1108/BPMJ-03-2014-0019

Alphabet (2018). Number of full-time Alphabet employees from 2007 to 2017. *Statista - the Statistics Portal*. Retrieved from https://www.statista.com/statistics/273744/number-of-full-time-google-employees/

Anand, G., Ward, P., Tatikonda, M., & Schilling, D. (2009). Dynamic capabilities through continuous improvement infrastructure. *Journal of Operations Management*, *27*(6), 444–461. https://doi.org/10.1016/j.jom.2009.02.002

Apple (2018). Anzahl der Mitarbeiter von Apple Inc. in den Geschäftsjahren 2005 bis 2017 (in 1.000). *Statista - the Statistics Portal*. Retrieved from https://de.statista.com/statistik/daten/studie/191540/umfrage/anzahl-der-mitarbeiter-von-apple-seit-2005/

Arnheiter, E. D., & Maleyeff, J. (2005). The integration of lean management and Six Sigma. *The TQM Magazine*, *17*(1), 5–18. https://doi.org/10.1108/09544780510573020

Baaij, M., Greeven, M., & van Dalen, J. (2004). Persistent Superior Economic Performance, Sustainable Competitive Advantage, and Schumpeterian Innovation. *European Management Journal*, *22*(5), 517–531. https://doi.org/10.1016/j.emj.2004.09.010

Banerjee, A., Bandyopadhyay, T., & Acharya, P. (2013). Data Analytics: Hyped Up Aspirations or True Potential? *Vikalpa: The Journal for Decision Makers*, *38*(4), 1–11.

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, *17*(1), 99–120. https://doi.org/10.1177/014920639101700108

Bell, S. (2006). *Lean enterprise systems: Using IT for continuous improvement*. *Wiley series in systems engineering and management*. Hoboken, NJ: Wiley. Retrieved from http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10305353 https://doi.org/10.1002/0471756466

Bell, S., & Orzen, M. (2010). *Lean IT: Enabling and Sustaining Your Lean Transformatation*. New York: Productivity Press.

Benninghaus, C., Elbe, C., Budde, L., & Friedli, T. (2018). *Digital Technologies – Evolution of production in high-wage countries*. Final Report. St. Gallen.

Bertagnolli, F. (2018). *Lean Management*. Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-13124-1

Bick, W. (2014). Warum Industrie 4.0 und Lean zwingend zusammengehören. *VDI-Z Integrierte Produktion*. (11), 47.

Bitkom (2018). Anzahl der Erwerbstätigen in der IT-Branche in Deutschland von 2007 bis 2018 nach Segment (in 1.000). *Statista - the Statistics Portal*. Retrieved from https://de.statista.com/statistik/daten/studie/186771/umfrage/erwerbstaetige-in-der-it-branche-in-deutschland/

Blum, M., & Schuh, G. (2017, April). Towards a Data-oriented Optimization of Manufacturing Processes. *Proceedings of the 19th International Conference on Enterprise Information Systems*, *Volume 8*, 257–264.

Bogner, A., Littig, B., & Menz, W. (2009). Introduction: Expert Interviews — An Introduction to a New Methodological Debate. In A. Bogner, B. Littig, & W. Menz (Eds.), *Interviewing Experts* (pp. 1–13). London: Palgrave Macmillan UK. https://doi.org/10.1057/9780230244276_1

Bokrantz, J., Skoogh, A., Berlin, C., & Stahre, J. (2017). Maintenance in digitalised manufacturing: Delphi-based scenarios for 2030. *International Journal of Production Economics*, *191*, 154–169. https://doi.org/10.1016/j.ijpe.2017.06.010

Brandt, P. (2017). *Industrie 4.0: Was lernen wir aus früheren Informatisierungswellen?* Düsseldorf.

Braun, M., Förster, U., & Vorspel-Rüter, F. (1988). *Mit CIM die Zukunft gestalten*. Frankfurt am Main.

Buer, S.-V., Strandhagen, J. O., & Chan, F. T. S. (2018). The link between Industry 4.0 and lean manufacturing: Mapping current research and establishing a research agenda. *International Journal of Production Research*, *56*(8), 2924–2940. https://doi.org/10.1080/00207543.2018.1442945

Cambridge University Press (2019a). Search term 'application'. Retrieved from https://dictionary.cambridge.org/dictionary/english/application

Cambridge University Press (2019b). Search term 'enabler'. Retrieved from https://dictionary.cambridge.org/de/worterbuch/englisch/enabler

Cambridge University Press (2019c). Search term 'real-time'. Retrieved from https://dictionary.cambridge.org/de/worterbuch/englisch/real-time

Card, D. (2017). The “black box” metaphor in machine learning. Retrieved from https://towardsdatascience.com/the-black-box-metaphor-in-machine-learning-4e57a3a1d2b0

Chen, J. C., & Chen, K.-M. (2014). Application of ORFPM system for lean implementation: an industrial case study. *The International Journal of Advanced Manufacturing Technology*, *72*(5), 839–852. https://doi.org/10.1007/s00170-014-5710-1

Chen, M., Mao, S., & Liu, Y. [Yunhao] (2014). Big Data: A Survey. *Mobile Networks and Applications*, *19*(2), 171–209. https://doi.org/10.1007/s11036-013-0489-0

Chen, Y. (2017). Integrated and Intelligent Manufacturing: Perspectives and Enablers. *Engineering*, *3*(5), 588–595. https://doi.org/10.1016/J.ENG.2017.04.009

Chiarini, A., Found, P., & Rich, N. [Nicholas] (Eds.) (2016). *Understanding the Lean Enterprise: Strategies, Methodologies, and Principles for a More Responsive Organization*. Cham: Springer International Publishing.

Choudhary, A. K., Harding, J. A., & Tiwari, M. K. (2009). Data mining in manufacturing: A review based on the kind of knowledge. *Journal of Intelligent Manufacturing*, *20*(5), 501–521. https://doi.org/10.1007/s10845-008-0145-x

Clegg, J. M. a. P. B., & Powell, D. (2013). ERP systems in lean production: New insights from a review of lean and ERP literature. *International Journal of Operations & Production Management*, *33*(11/12), 1490–1510. https://doi.org/10.1108/IJOPM-07-2010-0195

Cooper, H. M. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society*, *1*(1), 104–126. https://doi.org/10.1007/BF03177550

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4. ed.). Los Angeles: SAGE.

Davis, F. d. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319. https://doi.org/10.2307/249008

Dombrowski, U., Richter, T., & Krenkel, P. (2017). Interdependencies of Industrie 4.0 & Lean Production Systems: A Use Cases Analysis. *Procedia Manufacturing*, *11*, 1061–1068. https://doi.org/10.1016/j.promfg.2017.07.217

Duarte, S., & Cruz-Machado, V. (2013). Modelling lean and green: A review from business models. *International Journal of Lean Six Sigma*, *4*(3), 228–250. https://doi.org/10.1108/IJLSS-05-2013-0030

Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *Academy of Management Review*, *14*(4), 532–550. https://doi.org/10.5465/amr.1989.4308385

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory Building From Cases: Opportunities And Challenges. *Academy of Management Journal*, *50*(1), 25–32. https://doi.org/10.5465/amj.2007.24160888

Fleig, J. (2019). Wirtschaftlichkeitsrechnung und Investitionsrechnung. Retrieved from https://www.business-wissen.de/hb/wirtschaftlichkeitsrechnung-und-investitionsrechnung/

Flynn, B. B., Schroeder, R. G., & Flynn, E.J. (1999). World class manufacturing: an investigation of Hayes and Wheelwright's foundation. *Journal of Operations Management*, *17*(3), 249–269. https://doi.org/10.1016/S0272-6963(98)00050-3

Frankfurter Allgemeine Zeitung GmbH (2017). KÜNSTLICHE INTELLIGENZ: Ein Wettlauf wie im Kalten Krieg. Retrieved from https://www.faz.net/aktuell/wirtschaft/kuenstliche-intelligenz/kuenstliche-intelligenz-wer-gewinnt-den-wettlauf-usa-oder-china-15284415.html

Friedemann, M., Trapp, T. U., Stoldt, J., Langer, T., & Putz, M. (2016). A Framework for Information-driven Manufacturing. *Procedia CIRP*, *57*, 38–43. https://doi.org/10.1016/j.procir.2016.11.008

Friedli, T., Basu, P., Bellm, D., & Werani, J. (Eds.) (2013). *Leading Pharmaceutical Operational Excellence: Outstanding Practices and Cases*. Berlin, Heidelberg: Springer.

Friedli, T., & Bellm, D. (2013). OPEX: A Definition. In T. Friedli, P. Basu, D. Bellm, & J. Werani (Eds.), *Leading Pharmaceutical Operational Excellence: Outstanding Practices and Cases* (pp. 7–26). Berlin, Heidelberg: Springer.

Friedli, T., Lembke, N., Schneider, U. [Uli], & Gütter, S. (2013). The Current State of Operational Excellence Implementation: 10 Years of Benchmarking. In T. Friedli, P. Basu, D. Bellm, & J. Werani (Eds.), *Leading Pharmaceutical Operational*

Excellence: Outstanding Practices and Cases (pp. 35–58). Berlin, Heidelberg: Springer.

Friedli, T., Mänder, C., & Bellm, D. (2013). Matching Problems with Tools. In T. Friedli, P. Basu, D. Bellm, & J. Werani (Eds.), *Leading Pharmaceutical Operational Excellence: Outstanding Practices and Cases* (pp. 285–321). Berlin, Heidelberg: Springer.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Gassmann, O. (1999). Praxisnähe mit Fallstudienforschung. *Wissenschaftsmanagement*, *5*(3), 11–16.

Gerberich, T. (2011). *Lean oder MES in der Automobilzulieferindustrie: Ein Vorgehensmodell zur fallspezifischen Auswahl* (1. Aufl.). Wiesbaden: Gabler Verlag / Springer Fachmedien Wiesbaden GmbH Wiesbaden. https://doi.org/10.1007/978-3-8349-6754-1

Gerring, J. (2009). *Case study research: Principles and practices* (Reprinted.). Cambridge: Cambridge Univ. Press. Retrieved from http://www.loc.gov/catdir/enhancements/fy0703/2006051819-d.html

Gewohn, M., Usländer, T., Beyerer, J., & Sutschet, G. (2018). Digital Real-time Feedback of Quality-related Information to Inspection and Installation Areas of Vehicle Assembly. *Procedia CIRP*, *67*, 458–463. https://doi.org/10.1016/j.procir.2017.12.242

Ghobakhloo, M. (2018). The future of manufacturing industry: A strategic roadmap toward Industry 4.0. *Journal of Manufacturing Technology Management*, *29*(6), 910–936. https://doi.org/10.1108/JMTM-02-2018-0057

Hagerty, J. (2017). *Planning Guide for Data and Analytics*. Retrieved from https://www.gartner.com/binaries/content/assets/events/keywords/catalyst/catus8/2017_planning_guide_for_data_analytics.pdf

Handelsblatt GmbH (2019, June 7). KI-ENTWICKLUNG: Künstliche Intelligenz: Deutschland weit hinter den USA. Retrieved from https://www.handelsblatt.com/technik/vernetzt/ki-entwicklung-kuenstliche-intelligenz-deutschland-weit-hinter-den-usa/24434134.html

Harding, J. A., Shahbaz, M., Srinivas, & Kusiak, A. (2006). Data Mining in Manufacturing: A Review. *Journal of Manufacturing Science and Engineering*, *128*(4), 969. https://doi.org/10.1115/1.2194554

Hedelind, M., & Jackson, M. (2011). How to improve the use of industrial robots in lean manufacturing systems. *Journal of Manufacturing Technology Management*, *22*(7), 891–905. https://doi.org/10.1108/17410381111160951

Hicks, B. J. (2007). Lean information management: Understanding and eliminating waste. *International Journal of Information Management*, *27*(4), 233–249. https://doi.org/10.1016/j.ijinfomgt.2006.12.001

Hillmer, U. (2009). Existing Theories Considering Technology Adoption. In U. Hillmer (Ed.), *Technology Acceptance in Mechatronics: An Analysis of the Social and Personal Influences that affect Human Reaction to disruptive Technological Change in Mechatronics. Dissertation, University of East London 2009* (1st ed., pp. 9–28). Wiesbaden: Gabler Verlag. https://doi.org/10.1007/978-3-8349-8375-6_3

Hines, P., Holweg, M., & Rich, N. [Nick] (2004). Learning to evolve. *International Journal of Operations & Production Management*, *24*(10), 994–1011. https://doi.org/10.1108/01443570410558049

Hirmer, P., Wieland, M., Schwarz, H., Mitschang, B., Breitenbücher, U., Sáez, S. G., & Leymann, F. (2017). Situation recognition and handling based on executing situation templates and situation-aware workflows. *Computing*, *99*(2), 163–181. https://doi.org/10.1007/s00607-016-0522-9

Hirsch-Kreinsen, H., Hompel, M. ten, Ittermann, P., Dregger, J., Niehaus, J., Kirks, T., & Mättig, B. (2018). „Social Manufacturing and Logistics“ – Arbeit in der digitalisierten Produktion. In S. Wischmann & E. A. Hartmann (Eds.), *Zukunft der Arbeit – Eine praxisnahe Betrachtung* (Vol. 29, pp. 175–194). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-49266-6_13

Hitt, M. A., Xu, K., & Carnes, C. M. (2016). Resource based theory in operations management research. *Journal of Operations Management*, *41*, 77–94. https://doi.org/10.1016/j.jom.2015.11.002

Holweg, M. (2007). The genealogy of lean production. *Journal of Operations Management*, *25*(2), 420–437. https://doi.org/10.1016/j.jom.2006.04.001

Hopp, W. J., & Spearman, M. L. (2004). To Pull or Not to Pull: What Is the Question? *Manufacturing & Service Operations Management*, *6*(2), 133–148. https://doi.org/10.1287/msom.1030.0028

Illa, P. K., & Padhi, N. (2018). Practical Guide to Smart Factory Transition Using IoT, Big Data and Edge Analytics. *IEEE Access*, *6*, 55162–55170. https://doi.org/10.1109/ACCESS.2018.2872799

International Telecommunication Union (2012). Overview of the Internet of things: Internet of Things Global Standards Initiative. Retrieved from https://www.itu.int/en/ITU-T/gsi/iot/Pages/default.aspx

Jarupathirun, S., Ciganek, A. P., Chotiwankaewmanee, T., & Kerdpitak, C. (2009). Supply chain efficiencies through e-kanban: A case study. *International Journal of the Computer, the Internet and Management*, *17*(1), 55-1.

Kadane, S., & Bhatwadekar, S. (2011). Manufacturing Facility Layout Design and Optimization Using Simulation. *International Journal of Advanced Manufacturing Systems*, *2*.

Kagermann, H., Wahlster, W., & Helbig, J. (2013). *Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0: Abschlussbericht des Arbeitskreises Industrie 4.0.* Retrieved from https://www.bmbf.de/files/Umsetzungsempfehlungen_Industrie4_0.pdf

Kang, H. S., Lee, J. Y., Choi, S., Kim, H., Park, J. H., Son, J. Y., . . . Noh, S. D. (2016). Smart manufacturing: Past research, present findings, and future directions. *International Journal of Precision Engineering and Manufacturing-Green Technology*, *3*(1), 111–128. https://doi.org/10.1007/s40684-016-0015-5

Karre, H., Hammer, M., Kleindienst, M., & Ramsauer, C. (2017). Transition towards an Industry 4.0 State of the LeanLab at Graz University of Technology. *Procedia Manufacturing*, *9*, 206–213. https://doi.org/10.1016/j.promfg.2017.04.006

Ketokivi, M., & Choi, T. (2014). Renaissance of case research as a scientific method. *Journal of Operations Management*, *32*(5), 232–240. https://doi.org/10.1016/j.jom.2014.03.004

Kickuth, M., Friedli, T., Loch, C. H., & Chick, S. E. (2006). A Reference Model for Operational Excellence. In T. Friedli & B. Bastoen (Eds.), *Der Pharmazeutische Betrieb: Vol. 50. Operational excellence in the pharmaceutical industry*. Aulendorf: ECV - Ed.-Cantor-Verl.

Kieviet, A. (2016). Digitalisierung der Wertschöpfung: Auswirkung auf das Lean Management. In H. Künzel (Ed.), *Erfolgsfaktor Serie. Erfolgsfaktor Lean Management 2.0: Wettbewerbsfähige Verschlankung auf nachhaltige und kundenorientierte Weise* (pp. 41–59). Berlin, Heidelberg: Springer Gabler. https://doi.org/10.1007/978-3-662-49752-4_3

Kletti, J. (2015). *MES - Manufacturing Execution System*. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-46902-6

Kolberg, D., Knobloch, J., & Zühlke, D. (2016). Towards a lean automation interface for workstations. *International Journal of Production Research*, *55*(10), 2845–2856. https://doi.org/10.1080/00207543.2016.1223384

Kolberg, D., & Zühlke, D. (2015). Lean Automation enabled by Industry 4.0 Technologies. *IFAC-PapersOnLine*, *48*(3), 1870–1875. https://doi.org/10.1016/j.ifacol.2015.06.359

Krafcik, J. F. (1988). Triumph of the LEAN Production System. *Sloan Managemet Review*, *1988*(1), 41–52.

Kubicek, H. (1977). Heuristische Bezugsrahmen und heuristisch angelegte Forschungsdesign als Elemente einer Konstruktionsstrategie empirischer Forschung. In R. Köhler (Ed.), *Empirische und handlungstheoretische Forschungskonzeptionen in der Betriebswirtschaftslehre: Bericht über d. Tagung in Aachen, März 1976*. Stuttgart: Poeschel.

Kumar, R., Singh, S. P., & Lamba, K. (2018). Sustainable robust layout using Big Data approach: A key towards industry 4.0. *Journal of Cleaner Production*, *204*, 643–659. https://doi.org/10.1016/j.jclepro.2018.08.327

Künzel, H. (Ed.) (2016). *Erfolgsfaktor Serie*. *Erfolgsfaktor Lean Management 2.0: Wettbewerbsfähige Verschlankung auf nachhaltige und kundenorientierte Weise*. Berlin, Heidelberg: Springer Gabler. Retrieved from http://dx.doi.org/10.1007/978-3-662-49752-4 https://doi.org/10.1007/978-3-662-49752-4

Küpper, D., Heidemann, A., Ströhle, J., Spindelndreier, D., & Knizek, C. (2017). When Lean Meets Industry 4.0: The Next Level of Operational Excellence. Retrieved from http://image-src.bcg.com/Images/BCG-When-Lean-Meets-Industry-4.0-Dec-2017_tcm108-179091.pdf

Kusiak, A. (2017). Smart manufacturing must embrace big data. *Nature*, *544*(7648), 23–25. https://doi.org/10.1038/544023a

Kusiak, A. (2018). Smart manufacturing. *International Journal of Production Research*, *56*(1-2), 508–517. https://doi.org/10.1080/00207543.2017.1351644

Lander, E., & Liker, J. K. (2007). The Toyota Production System and art: Making highly customized and creative products the Toyota way. *International Journal of Production Research*, *45*(16), 3681–3698. https://doi.org/10.1080/00207540701223519

Landherr, M., Schneider, U. [Ulrich], & Bauernhansl, T. (2016). The Application Center Industrie 4.0 - Industry-driven Manufacturing, Research and Development. *Procedia CIRP*, *57*, 26–31. https://doi.org/10.1016/j.procir.2016.11.006

Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. *Business & Information Systems Engineering*, *6*(4), 239–242. https://doi.org/10.1007/s12599-014-0334-4

Lee, J. [Jay], Bagheri, B., & Kao, H.-a. (2015). A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, *3*, 18–23. https://doi.org/10.1016/j.mfglet.2014.12.001

Lee, J. [Jay], Kao, H.-a., & Yang, S. (2014). Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment. *Procedia CIRP*, *16*, 3–8. https://doi.org/10.1016/j.procir.2014.02.001

Lee, J. [Jay], Lapira, E., Bagheri, B., & Kao, H.-a. (2013). Recent advances and trends in predictive manufacturing systems in big data environment. *Manufacturing Letters*, *1*(1), 38–41. https://doi.org/10.1016/j.mfglet.2013.09.005

Lenz, J., Kotschenreuther, J., & Westkaemper, E. (2017). Energy Efficiency in Machine Tool Operation by Online Energy Monitoring Capturing and Analysis. *Procedia CIRP*, *61.* https://doi.org/10.1016/j.procir.2016.11.202

Levy, Y., & Ellis, T. (2006). A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research. *Informing Science: The International Journal of an Emerging Transdiscipline*, *9*, 181–212. https://doi.org/10.28945/479

Lewis, M. A. (2000). Lean production and sustainable competitive advantage. *International Journal of Operations & Production Management*, *20*(8), 959–978. https://doi.org/10.1108/01443570010332971

Li, J., Tao, F., Cheng, Y., & Zhao, L. (2015). Big Data in product lifecycle management. *The International Journal of Advanced Manufacturing Technology*, *81*(1), 667–684. https://doi.org/10.1007/s00170-015-7151-x

Li, X., Di Li, Wan, J., Vasilakos, A. V., Lai, C.-F., & Wang, S. (2017). A review of industrial wireless networks in the context of Industry 4.0. *Wireless Networks*, *23*(1), 23–41. https://doi.org/10.1007/s11276-015-1133-7

Li, Z., Wang, Y. [Yi], & Wang, K.-S. (2017). Intelligent predictive maintenance for fault diagnosis and prognosis in machine centers: Industry 4.0 scenario. *Advances in Manufacturing*, *5*(4), 377–387. https://doi.org/10.1007/s40436-017-0203-8

Liao, W., & Wang, Y. [Ying] (2013). Data-driven Machinery Prognostics Approach using in a Predictive Maintenance Model. *Journal of Computers*, *8*(1). https://doi.org/10.4304/jcp.8.1.225-231

Lidong, W., & Guanghui, W. (2016). Big Data in Cyber-Physical Systems, Digital Manufacturing and Industry 4.0. *International Journal of Engineering and Manufacturing*, *6*(4), 1–8. https://doi.org/10.5815/ijem.2016.04.01

Liker, J. K. (2004). *The Toyota way: 14 management principles from the world's greatest manufacturer*. New York, NY: McGraw-Hill. Retrieved from http://www.loc.gov/catdir/bios/mh041/2004300007.html

Liker, J. K., & Morgan, J. M. (2006). The Toyota Way in Services: The Case of Lean Product Development. *Academy of Management Perspectives*, *20*(2), 5–20. https://doi.org/10.5465/amp.2006.20591002

Lodewijks, G., Kraijema, S., Godjevac, M., & Corman, F. Optimizing preventive maintenance policy: A data-driven application for a light rail braking system. *Proceedings of the Institution of Mechanical Engineers.* Advance online publication. https://doi.org/10.3929/ethz-b-000182703

Lokrantz, A., Gustavsson, E., & Jirstrand, M. (2018). Root cause analysis of failures and quality deviations in manufacturing using machine learning. *Procedia CIRP*, *72*, 1057–1062. https://doi.org/10.1016/j.procir.2018.03.229

Lorenz, R., Buess, P., Macuvele, J., Friedli, T., & Netland, T. H. (2019). Lean and Digitalization—Contradictions or Complements? In F. Ameri, K. E. Stecke, G. von Cieminski, & D. Kiritsis (Eds.), *IFIP Advances in Information and Communication Technology: Vol. 566. Advances in production management systems: Production management for the factory of the future : IFIP WG 5.7 International Conference, APMS 2019, Austin, TX, USA, September 1-5, 2019 : proceedings* (Vol. 566, pp. 77–84). Cham: Springer. https://doi.org/10.1007/978-3-030-30000-5_10

Louw, L., & Walker, M. (2018). Design and implementation of a low cost RFID track and trace system in a learning factory. *Procedia Manufacturing*, *23*, 255–260. https://doi.org/10.1016/j.promfg.2018.04.026

Lu, Y. (2017). Industry 4.0: A survey on technologies, applications and open research issues. *Journal of Industrial Information Integration*, *6*, 1–10. https://doi.org/10.1016/j.jii.2017.04.005

Macuvele, J., Buess, P., & Friedli, T. (2018). *Lean2020 - The Future of Operational Excellence: Final Report*. St. Gallen.

Maguire, K. (2016). Lean and IT—Working Together? An Exploratory Study of the Potential Conflicts Between Lean Thinking and the Use of Information Technology in Organisations Today. In A. Chiarini, P. Found, & N. Rich (Eds.), *Understanding the Lean Enterprise: Strategies, Methodologies, and Principles for a More Responsive Organization* (pp. 31–60). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-19995-5_2

Majstorovic, V., Stojadinovic, S., Jakovljevic, Z., Zivkovic, S., Djurdjanovic, D., Kostic, J., & Gligorijevic, N. (2018). Cyber-Physical Manufacturing Metrology Model (CPM3) – Big Data Analytics Issue. *Procedia CIRP*, *72*, 503–508. https://doi.org/10.1016/j.procir.2018.03.091

Manyika, J. [J.], Chui, M. [M.], Brown, B., Bughin, J. [J.], Dobbs, R. [R.], Roxburgh, C., & Hung Byers, A. (2011). Big data: The next frontier for innovation, competition, and productivity. Retrieved from https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_full_report.ashx

Manyika, J. [James], Chui, M. [Michael], Bisson, P., Woetzel, J., Dobbs, R. [Richard], Bughin, J. [Jacques], & Aharon, D. (June 2015). *The Internet of Things: Mapping the value beyond the hype: Executive Summary*. Retrieved from https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/The%20Internet%20of%20Things%20The%20value%20of%20digitizing%20the%20physical%20world/Unlocking_the_potential_of_the_Internet_of_Things_Executive_summary.ashx

Markarian, J. (2018). Modernizing Pharma Manufacturing. *Pharmaceutical Technology Europe*, *30*(4), 10–12,35. Retrieved from https://search.proquest.com/docview/2056435254?accountid=28962

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*(4), 370–396. https://doi.org/10.1037/h0054346

Matyas, K., Nemeth, T., Kovacs, K., & Glawar, R. (2017). A procedural approach for realizing prescriptive maintenance planning in manufacturing industries. *CIRP Annals*, *66*(1), 461–464. https://doi.org/10.1016/j.cirp.2017.04.007

Mayr, A., Weigelt, M., Kühl, A., Grimm, S., Erll, A., Potzel, M., & Franke, J. (2018). Lean 4.0 - A conceptual conjunction of lean management and Industry 4.0. *Procedia CIRP*, *72*, 622–628. https://doi.org/10.1016/j.procir.2018.03.292

Mehta, P., Butkewitsch-Choze, S., & Seaman, C. (2018). Smart manufacturing analytics application for semi-continuous manufacturing process – a use case. *Procedia Manufacturing*, *26*, 1041–1052. https://doi.org/10.1016/j.promfg.2018.07.138

Meissner, A., Müller, M., Hermann, A., & Metternich, J. (2018). Digitalization as a catalyst for lean production: A learning factory approach for digital shop floor management. *Procedia Manufacturing*, *23*, 81–86. https://doi.org/10.1016/j.promfg.2018.03.165

Meredith, J. (1998). Building operations management theory through case and field research. *Journal of Operations Management*, *16*(4), 441–454. https://doi.org/10.1016/S0272-6963(98)00023-0

Metternich, J., Müller, M., Meudt, T., & Schaede, C. (2017). Lean 4.0 – zwischen Widerspruch und Vision. *ZWF*, *112*, 346–348.

Meudt, T., Metternich, J., & Abele, E. (2017). Value stream mapping 4.0: Holistic examination of value stream and information logistics in production. *CIRP Annals*, *66*(1), 413–416. https://doi.org/10.1016/j.cirp.2017.04.005

Meziane, F., Vadera, S., Kobbacy, K., & Proudlove, N. (2000). Intelligent systems in manufacturing: current developments and future prospects. *Integrated Manufacturing Systems*, *11*(4), 218–238. https://doi.org/10.1108/09576060010326221

Microsoft (2018). Anzahl der Mitarbeiter der Microsoft Corporation weltweit in den Geschäftsjahren 2005 bis 2017 (in 1.000). *Statista - the Statistics Portal*. Retrieved from https://de.statista.com/statistik/daten/studie/191876/umfrage/anzahl-der-mitarbeiter-der-microsoft-corporation-seit-2005/

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2. ed [Nachdr.]). Thousand Oaks, Calif.: SAGE.

Mittal, S., Khan, M. A., Romero, D., & Wuest, T. (2017). Smart manufacturing: Characteristics, technologies and enabling factors. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 0954405417736547. https://doi.org/10.1177/0954405417736547

Moore, G. A. (1995). *Insight the tornado:: Marketing strategies from Silicon Valley's cutting edge*. New York: HarperCollins Publishers.

Mrugalska, B., & Wyrwicka, M. K. (2017). Towards Lean Production in Industry 4.0. *Procedia Engineering*, *182*, 466–473. https://doi.org/10.1016/j.proeng.2017.03.135

Munirathinam, S., & Ramadoss, B. (2014). Big data predictive analtyics for proactive semiconductor equipment maintenance. In J. Lin (Ed.), *Ieee International Conference on Big Data (Big Data), 2014: 27 - 30 Oct. 2014, Washington, DC, USA* (pp. 893–902). Piscataway, NJ: IEEE. https://doi.org/10.1109/BigData.2014.7004320

National Institute of Standards and Technology (2017). Smart Manufacturing Operations Planning and Control Program. Retrieved from https://www.nist.gov/programs-projects/smart-manufacturing-operations-planning-and-control-program

Nemeth, T., Ansari, F., Sihn, W., Haslhofer, B., & Schindler, A. (2018). PriMa-X: A reference model for realizing prescriptive maintenance and assessing its maturity enhanced by machine learning. *Procedia CIRP*, *72*, 1039–1044. https://doi.org/10.1016/j.procir.2018.03.280

Netland, T. H. (2015). Industry 4.0: Where does it leave lean? *Lean Management Journal*, *5*(2), 22–23.

Ngo, Q. H., & Schmitt, R. H. (2016). A Data-based Approach for Quality Regulation. *Procedia CIRP*, *57*, 498–503. https://doi.org/10.1016/j.procir.2016.11.086

O'Donovan, P., Leahy, K., Bruton, K., & O'Sullivan, D. T. J. (2015a). Big data in manufacturing: A systematic mapping study. *Journal of Big Data*, *2*(1), 20. https://doi.org/10.1186/s40537-015-0028-x

O'Donovan, P., Leahy, K., Bruton, K., & O'Sullivan, D. T. J. (2015b). An industrial big data pipeline for data-driven analytics maintenance applications in large-scale smart manufacturing facilities. *Journal of Big Data*, *2*(1), 25. https://doi.org/10.1186/s40537-015-0034-z

Ohno, T. (1988). *Toyota production system: Beyond large-scale production*. New York: Productivity Press.

Oliff, H., & Liu, Y. [Ying] (2017). Towards Industry 4.0 Utilizing Data-Mining Techniques: A Case Study on Quality Improvement. *Procedia CIRP*, *63*, 167–172. https://doi.org/10.1016/j.procir.2017.03.311

Peinado, J., Graeml, A. R., & Vianna, F. (2018). Operations management body of knowledge and its relevance to manufacturing and service organizations. *Revista De Gestão*, *25*(4), 373–389. https://doi.org/10.1108/REGE-03-2018-0049

Pettersen, J. (2009). Defining lean production: Some conceptual and practical issues. *The TQM Journal*, *21*(2), 127–142. https://doi.org/10.1108/17542730910938137

Pettigrew, A. M. (1990). Longitudinal Field Research on Change: Theory and Practice. *Organization Science*, *1*(3), 267–292. https://doi.org/10.1287/orsc.1.3.267

Pham, D. T., & Afify, A. A. (2005). Machine-learning techniques and their applications in manufacturing. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *219*(5), 395–412. https://doi.org/10.1243/095440505X32274

Pilloni, V. (2018). How Data Will Transform Industrial Processes: Crowdsensing, Crowdsourcing and Big Data as Pillars of Industry 4.0. *Future Internet*, *10*(3), 24. https://doi.org/10.3390/fi10030024

Powell, D., Alfnes, E., Strandhagen, J. O., & Dreyer, H. (2013). The concurrent application of lean production and ERP: Towards an ERP-based lean implementation process. *Computers in Industry*, *64*(3), 324–335. https://doi.org/10.1016/j.compind.2012.12.002

Prinz, C., Kreggenfeld, N., & Kuhlenkötter, B. (2018). Lean meets Industrie 4.0 – a practical approach to interlink the method world and cyber-physical world. *Procedia Manufacturing*, *23*, 21–26. https://doi.org/10.1016/j.promfg.2018.03.155

Provost, F., & Fawcett, T. (2013a). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*, *1*(1), 51–59. https://doi.org/10.1089/big.2013.1508

Provost, F., & Fawcett, T. (2013b). *Data science for business: What you need to know about data mining and data-analytic thinking* (1st ed.). Sebastopol, CA: O'Reilly Media.

Qi, Q., & Tao, F. (2018). Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison. *IEEE Access*, 1. https://doi.org/10.1109/ACCESS.2018.2793265

Qin, J., Liu, Y. [Ying], & Grosvenor, R. (2016). A Categorical Framework of Manufacturing for Industry 4.0 and Beyond. *Procedia CIRP*, *52*, 173–178. https://doi.org/10.1016/j.procir.2016.08.005

Rafique, M. Z., Ab Rahman, M. N., Saibani, N., Arsad, N., & Saadat, W. (2016). RFID impacts on barriers affecting lean manufacturing. *Industrial Management & Data Systems*, *116*(8), 1585–1616. https://doi.org/10.1108/IMDS-10-2015-0427

Randolph, & Justus (2009). A Guide to Writing the Dissertation Literature Review - Practical Assessment, Research & Evaluation. *Practical Assessment, Research & Evaluation*, *14*(13). Retrieved from http://pareonline.net/getvn.asp?v=14&n=13

Ray Y Zhong, Xun Xu, & Lihui Wang (2017). IoT-enabled Smart Factory Visibility and Traceability Using Laser-scanners. *Procedia Manufacturing*, *10*, 1–14. https://doi.org/10.1016/j.promfg.2017.07.103

Riezebos, J., & Klingenberg, W. (2009). Advancing lean manufacturing, the role of IT. *Computers in Industry*, *60*(4), 235–236. https://doi.org/10.1016/j.compind.2009.01.005

Rimpault, X., Balazinski, M., & Chatelain, J.-F. (2018). Fractal Analysis Application Outlook for Improving Process Monitoring and Machine Maintenance in Manufacturing 4.0. *Journal of Manufacturing and Materials Processing*, *2*(3), 62. https://doi.org/10.3390/jmmp2030062

Rogers, E. M. (1983). *Diffusion of innovations* (3. ed.). New York, NY: Free Press.

Rossini, M., Costa, F., Tortorella, G., & Portioli-Staudacher, A. (2019). The interrelation between Industry 4.0 and lean production: an empirical study on European manufacturers. *The International Journal of Advanced Manufacturing Technology*. Retrieved from https://doi.org/10.1007/s00170-019-03441-7

Russom, P. (2015). *Bringing Modern Data-Driven Applications to the Enterprise*. Retrieved from TDWI Research website: https://tdwi.org/research/2015/03/checklist-bringing-modern-data-driven-applications-to-the-enterprise.aspx

Rüttimann, B. G., & Stöckli, M. (2016). Lean and Industry 4.0—Twins, Partners, or Contenders? A Due Clarification Regarding the Supposed Clash of Two Production Systems. *Journal of Service Science and Management*, *09*, 485. https://doi.org/10.4236/jssm.2016.96051

Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, *3*(3), 210–229. https://doi.org/10.1147/rd.33.0210

Sanders, A., Elangeswaran, C., & Wulfsberg, J. (2016). Industry 4.0 implies lean manufacturing: Research activities in industry 4.0 function as enablers for lean manufacturing. *Journal of Industrial Engineering and Management*, *9*(3), 811. https://doi.org/10.3926/jiem.1940

Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research methods for business students* (5. ed.). Harlow: Financial Times Prentice Hall.

Saygin, C., & Sarangapani, J. (2006). Rfid on the Manufacturing Shop Floor: Applications and Challenges. *IIE Annual Conference. Proceedings*, 1–6.

Schöning, H., & Dorchain, M. (2014). Data Mining und Analyse. In T. Bauernhansl, M. ten Hompel, & B. Vogel-Heuser (Eds.), *Industrie 4.0 in Produktion, Automatisierung und Logistik: Anwendung · Technologien · Migration* (pp. 543–554). Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-04682-8_27

Segura Velandia, D. M., Kaur, N., Whittow, W. G., Conway, P. P., & West, A. A. (2016). Towards industrial internet of things: Crankshaft monitoring, traceability and tracking using RFID. *Robotics and Computer-Integrated Manufacturing*, *41*, 66–77. https://doi.org/10.1016/j.rcim.2016.02.004

Shah, R., & Ward, P. T. (2003). Lean manufacturing: Context, practice bundles, and performance. *Journal of Operations Management*, *21*(2), 129–149. https://doi.org/10.1016/S0272-6963(02)00108-0

Shah, R., & Ward, P. T. (2007). Defining and developing measures of lean production. *Journal of Operations Management*, *25*(4), 785–805. https://doi.org/10.1016/j.jom.2007.01.019

Shao, G., Shin, S.-J., & Jain, S. (2014). Data analytics using simulation for smart manufacturing. In A. Tolk (Ed.), *Winter Simulation Conference (WSC), 2014: 7 - 10 Dec. 2014, Savannah, GA* (pp. 2192–2203). Piscataway, NJ: IEEE. https://doi.org/10.1109/WSC.2014.7020063

Shuradze, G., & Wagner, H.-T. (2016). Towards a Conceptualization of Data Analytics Capabilities. In *2016 49th Hawaii International Conference on System Sciences (HICSS)* (pp. 5052–5064). IEEE. https://doi.org/10.1109/HICSS.2016.626

Sipsas, K., Alexopoulos, K., Xanthakis, V., & Chryssolouris, G. (2016). Collaborative Maintenance in flow-line Manufacturing Environments: An Industry 4.0 Approach. *Procedia CIRP*, *55*, 236–241. https://doi.org/10.1016/j.procir.2016.09.013

Sokovic, M., Pavletic, D., & Kern Pipan, K. (2010). Quality Improvement Methodologies – PDCA Cycle, RADAR Matrix, DMAIC and DFSS. *Journal of Achievements in Materials and Manufacturing Engineering*, *43*(1), 476–483.

Sommer, L. (2015). Industrial revolution - industry 4.0: Are German manufacturing SMEs the first victims of this revolution? *Journal of Industrial Engineering and Management*, *8*(5). https://doi.org/10.3926/jiem.1470

Sony, M. (2018). Industry 4.0 and lean management: A proposed integration model and research propositions. *Production & Manufacturing Research*, *6*(1), 416–432. https://doi.org/10.1080/21693277.2018.1540949

Spath, D., Ganschar, O., Gerlach, S., Hämmerle, M., Krause, T., & Schlund, S. (2013). *duktionsarbeit der Zukunft - Industrie 4.0*. Stuttgart. Retrieved from https://www.iao.fraunhofer.de/images/iao-news/produktionsarbeit-der-zukunft.pdf

Statista (2018). Größte Automobilhersteller weltweit nach Fahrzeugabsatz im Jahr 2017 (in Millionen). *Statista - the Statistics Portal*. Retrieved from https://de.statista.com/statistik/daten/studie/173795/umfrage/automobilhersteller-nach-weltweitem-fahrzeugabsatz/

Statistisches Bundesamt (2018). Anzahl der Beschäftigten im Verarbeitenden Gewerbe in Deutschland von April 2017 bis April 2018. *Statista - the Statistics Portal*. Retrieved from https://de.statista.com/statistik/daten/studie/161054/umfrage/beschaeftigte-verarbeitendes-gewerbe/

Taherdoost, H. (2018). A review of technology acceptance and adoption models and theories. *Procedia Manufacturing*, *22*, 960–967. https://doi.org/10.1016/j.promfg.2018.03.137

Taisch, M., Arena, D. N., Gorobtcova, P., Kiritsis, D., Luglietti, R., May, G., . . . Wuest, T. (2018). *The 2018 World Manufacturing Forum Report: Recommendations for the Future of Manufacturing*. Retrieved from World Manufacturing Forum website: https://www.worldmanufacturingforum.org/report

Tantik, E., & Anderl, R. (2016). Industrie 4.0: Using Cyber-physical Systems for Value-stream Based Production Evaluation. *Procedia CIRP*, *57*, 207–212. https://doi.org/10.1016/j.procir.2016.11.036

Tao, F., Cheng, Y., Xu, L. D., Zhang, L., & Li, B. H. (2014). CCIoT-CMfg: Cloud Computing and Internet of Things-Based Cloud Manufacturing Service System. *IEEE Transactions on Industrial Informatics*, *10*(2), 1435–1442. https://doi.org/10.1109/TII.2014.2306383

Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. *Journal of Manufacturing Systems.* Advance online publication. https://doi.org/10.1016/j.jmsy.2018.01.006

Thoben, K.-D., Wiesner, S., & Wuest, T. (2017). “Industrie 4.0” and Smart Manufacturing – A Review of Research Issues and Application Examples. *International Journal of Automation Technology*, *11*(1), 4–16. https://doi.org/10.20965/ijat.2017.p0004

Tomczak, T. (1992). Forschungsmethoden in der Marketingwissenschaft. Ein Plädoyer für den qualitativen Forschungsansatz. *Marketing ZFP*, *14*(2), 77–87. https://doi.org/10.15358/0344-1369-1992-2-77

Torraco, R. J. (2005). Writing Integrative Literature Reviews: Guidelines and Examples. *Human Resource Development Review*, *4*(3), 356–367. https://doi.org/10.1177/1534484305278283

Tortorella, G., & Fettermann, D. (2018). Implementation of Industry 4.0 and lean production in Brazilian manufacturing companies. *International Journal of Production Research*, *56*(8), 2975–2987. https://doi.org/10.1080/00207543.2017.1391420

Towill, D. R. (2007). Exploiting the DNA of the Toyota Production System. *International Journal of Production Research*, *45*(16), 3619–3637. https://doi.org/10.1080/00207540701223436

Toyota Material Handling (n.d.). Toyota Production System: and what it means for business. Retrieved from http://www.bt-forklifts.com/SiteCollectionDocuments/PDF%20files/Toyota%20Production%20System%20Brochure.pdf

Trinczek, R. (2009). How to Interview Managers? Methodical and Methodological Aspects of Expert Interviews as a Qualitative Method in Empirical Social Research. In A. Bogner, B. Littig, & W. Menz (Eds.), *Interviewing Experts* (pp. 203–216). London: Palgrave Macmillan UK. https://doi.org/10.1057/9780230244276_10

Ulrich, H. (1982). Anwendungsorientierte Wissenschaft. *Die Unternehmung*, *36*(1), 1–10.

Ulrich, H. (1984). Die Betriebswirtschaftslehre als anwendungsorientierte Sozialwissenschaft. In H. Ulrich, D. Thomas, & P. Gilbert (Eds.), *Management* (pp. 168–199). Bern: Haupt.

Uriarte, A. G., Ng, A. H.C., & Moris, M. U. (2018). Supporting the lean journey with simulation and optimization in the context of Industry 4.0. *Procedia Manufacturing*, *25*, 586–593. https://doi.org/10.1016/j.promfg.2018.06.097

Vallhagen, J., Almgren, T., & Thörnblad, K. (2017). Advanced use of Data as an Enabler for Adaptive Production Control using Mathematical Optimization – An Application of Industry 4.0 Principles. *Procedia Manufacturing*, *11*, 663–670. https://doi.org/10.1016/j.promfg.2017.07.165

Van Amelsvoort, P., & Benders, J. (1996). Team time: a model for developing self-directed work teams. *International Journal of Operations & Production Management*, *16*(2), 159–170. https://doi.org/10.1108/01443579610109901

Venkatesh, Morris, & Davis (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425. https://doi.org/10.2307/30036540

Vogel-Heuser, B., Bauernhansl, T., & Hompel, M. ten (Eds.) (2017). *Springer Reference Technik*. *Handbuch Industrie 4.0: Bd. 4: Allgemeine Grundlagen* (2. Auflage). Berlin: Springer Vieweg. https://doi.org/10.1007/978-3-662-53254-6

Volkswagen (2018). Beschäftigtenzahl der Volkswagen AG in den Jahren 2013 bis 2017 (jeweils am Jahresende). *Statista - the Statistics Portal*. Retrieved from https://de.statista.com/statistik/daten/studie/181504/umfrage/beschaeftigtenzahl-der-volkswagen-ag/

Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., & Platfaut, R. (2015). Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. *Communications of the Association for Information Systems*, *37*, 205–224.

Voss, C., Tsikriktsis, N., & Frohlich, M. (2002). Case research in operations management. *International Journal of Operations & Production Management*, *22*(2), 195–219. https://doi.org/10.1108/01443570210414329

Wagner, T., Herrmann, C., & Thiede, S. (2017). Industry 4.0 Impacts on Lean Production Systems. *Procedia CIRP*, *63*, 125–131. https://doi.org/10.1016/j.procir.2017.02.041

Waibel, M. W., Steenkamp, L. P., Moloko, N., & Oosthuizen, G. A. (2017). Investigating the Effects of Smart Production Systems on Sustainability Elements. *Procedia Manufacturing*, *8*, 731–737. https://doi.org/10.1016/j.promfg.2017.02.094

Wan, J., Yang, J., Wang, Z., & Hua, Q. (2018). Artificial Intelligence for Cloud-Assisted Smart Factory. *IEEE Access*, *6*, 55419–55430. https://doi.org/10.1109/ACCESS.2018.2871724

Wang, B., Zhao, J.-y., Wan, Z.-g., Ma, J.-h. [Ji-hong], LI, H., & Ma, J. [Jian] (2016). Lean Intelligent Production System and Value Stream Practice. *DEStech Transactions on Economics and Management.* Advance online publication. https://doi.org/10.12783/dtem/icem2016/4106

Wang, S., Wan, J., Li, D., & Zhang, C. (2016). Implementing Smart Factory of Industrie 4.0: An Outlook. *International Journal of Distributed Sensor Networks*, *12*(1), 3159805. https://doi.org/10.1155/2016/3159805

Ward, P., & Zhou, H. (2006). Impact of Information Technology Integration and Lean/Just-In-Time Practices on Lead-Time Performance *Decision Sciences*, *37*(2), 177–203. https://doi.org/10.1111/j.1540-5915.2006.00121.x

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, *26*(2), xiii.

Wenking, M., Benninghaus, C., & Groggert, S. (2017). Die Zukunft von Manufacturing Data Analytics Implikationen für eine erfolgreiche Datennutzung im produzierenden Umfeld. *Industrie Management*. (4), 33–37. Retrieved from https://www.wiso-net.de/document/IM__BBE6FCC66E27D2FB130F473E5BD95617

Wolf, J. (2011). *Organisation, Management, Unternehmensführung: Theorien, Praxisbeispiele und Kritik* (4.th ed.). *Gabler-Lehrbuch*. Wiesbaden: Gabler.

Womack, J. P., & Jones, D. T. (2003). *Lean thinking: Banish waste and create wealth in your corporation* (Rev. and updated ed.). New York: Free Press.

Womack, J. P., Jones, D. T., & Roos, D. (1990). *The machine that changed the world: Based on the Massachusetts Institute of Technology 5 million dollar 5 year study on the future of the automobile*. New York, NY: Rawson.

Wuest, T., Irgens, C., & Thoben, K.-D. (2014). An approach to monitoring quality in manufacturing using supervised machine learning on product state data. *Journal of Intelligent Manufacturing*, *25*(5), 1167–1180. https://doi.org/10.1007/s10845-013-0761-y

Wuest, T., Weimer, D., Irgens, C., & Thoben, K.-D. (2016). Machine learning in manufacturing: Advantages, challenges, and applications. *Production & Manufacturing Research*, *4*(1), 23–45. https://doi.org/10.1080/21693277.2016.1192517

Yan, J., Meng, Y., Lu, L., & Li, L. (2017). Industrial Big Data in an Industry 4.0 Environment: Challenges, Schemes, and Applications for Predictive Maintenance. *IEEE Access*, *5*, 23484–23491. https://doi.org/10.1109/ACCESS.2017.2765544

Yin, R. (2009). *Case study research: Design and methods* (4. ed.). *Applied social research methods series: Vol. 5*. Los Angeles: SAGE.

Yin, S., & Kaynak, O. (2015). Big Data for Modern Industry: Challenges and Trends [Point of View]. *Proceedings of the IEEE*, *103*(2), 143–146. https://doi.org/10.1109/JPROC.2015.2388958

Yunusa-kaltungo, A., & Sinha, J. K. (2017). Effective vibration-based condition monitoring (eVCM) of rotating machines. *Journal of Quality in Maintenance Engineering*, *23*(3), 279–296. Retrieved from https://search.proquest.com/docview/1933250984?accountid=28962

ZEIT ONLINE GmbH (2018, November 28). Künstliche Intelligenz: Jetzt mal richtig. Retrieved from https://www.zeit.de/2018/48/kuenstliche-intelligenz-foerderung-geld-investition-ideen

Zhang, Y., Ren, S., Liu, Y. [Yang], & Si, S. (2017). A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products. *Journal of Cleaner Production*, *142*, 626–641. https://doi.org/10.1016/j.jclepro.2016.07.123

Zhang, Y., Zhang, G., Wang, J., Sun, S., Si, S., & Yang, T. (2014). Real-time information capturing and integration framework of the internet of manufacturing things. *International Journal of Computer Integrated Manufacturing*, *28*(8), 811–822. https://doi.org/10.1080/0951192X.2014.900874

Zhong, R. Y., Huang, G. Q., Dai, Q. Y., & Zhang, T. (2014). Mining SOTs and dispatching rules from RFID-enabled real-time shopfloor production data. *Journal of Intelligent Manufacturing*, *25*(4), 825–843. https://doi.org/10.1007/s10845-012-0721-y

Zhu, X., Qiao, F., & Cao, Q. (2017). Industrial big data–based scheduling modeling framework for complex manufacturing system. *Advances in Mechanical Engineering*, *9.* https://doi.org/10.1177/1687814017726289

Appendix

Appendix A: List of journals identified in the literature review and used in this research

No.	Journal	Database
1	Applied Sciences	Web of Science
2	Business Research	Proquest
3	Central European Business Review	Proquest
4	Computers & Industrial Engineering	Web of Science
5	Computing	EBSCOhost
6	Engineering	Web of Science
7	Enterprise Information Systems	EBSCOhost
8	IEEE Access	Web of Science
9	Independent Journal of Management & Production	Proquest
10	Industrial Management & Data Systems	Proquest
11	International Journal of Distributed Sensor Networks	Web of Science
12	International Journal of Information, Business and Management	Proquest
13	International Journal of Logistics Management	Proquest
14	International Journal of Operations & Production Management	Emarald Insight
15	International Journal of Production Research	Web of Science
16	Journal of Big Data	Proquest
17	Journal of Cleaner Production	EBSCOhost
18	Journal of Economic & Management Perspectives	Proquest
19	Journal of Enterprise Information Managemen	Emarald Insight
20	Journal of Industrial Engineering and Management	Web of Science
21	Journal of Innovation Management	Proquest
22	Journal of Intelligent Manufacturing	Proquest
23	Journal of International Trade, Logistics and Law	Proquest
24	Journal of Manufacturing Systems	Web of Science
25	Journal of Manufacturing Technology Management	Emarald Insight
26	Journal of Operations Management	Proquest
27	Journal of Quality in Maintenance Engineering	Proquest

No.	Journal	Database
28	LogForum / Journal of Logistics	Proquest
29	Logistics & Transport	EBSCOhost
30	Materials	Web of Science
31	Pharmaceutical Technology Europe	Proquest
32	Production & Manufacturing Research	Web of Science
33	Production and operations management	EBSCOhost
34	Research Technology Management	EBSCOhost
35	Technological & Economic Development of Economy	EBSCOhost
36	The Int. Journal of Advanced Manufacturing Technology	Web of Science
37	Wireless Networks	Proquest

Appendix B: Results of systematic literature review

Search terms and oeprators	Proquest	EmeraldInsight	EBSCOhost	Web of Science	ScienceDirect
1) noft(Lean OR TPS OR Toyota Production System) AND noft(production OR manufactur* OR technology) AND noft(practice* OR "")	209[1] (6)	449 (2)	305 (1)	38 (4)	59 (7)
2) (noft(Smart OR Digital) AND noft(factory OR manufactur* OR technology) OR noft (Industry 4.0 OR Industrie 4.0)) AND noft(data OR "")	264[2] (20)	200 (3)	610 (9)	193 (6)	225 (34)
3) noft(data) AND noft(analytics OR analysis OR mining OR visualization OR optimization OR "") AND noft(factory OR manufactur* OR production OR techniques OR methods)	848[3] (8)	273 (2)	195 (5)	116 (4)	19 (0)
Number of total hits and number of articles considered relevant after scanning abstract, in brackets.					
Comment Proquest Only relevant journals selected, limited to full text access and peer reviewed scholarly articles, only English and German, search period 01.01.2008 - 31.12.2018. , keyword search in abstract In Noft = not in full text. Search link (valid until 01.10.2020): Link [1], Link[2], Link[3]					
Comment EmeraldInsight Limited to full text access and peer reviewed scholarly articles, only English and German, search period 01.01.2008 - 31.12.2018., keyword search in abstract Search 2: search term "data" in "title",					
Comment EBSCOhost Only relevant journals selected, limited to full text access and peer reviewed scholarly articles, only English and German, search period 01.01.2008 - 31.12.2018. keyword search in abstract, Search 1: search term (Lean OR TPS OR Toyota Production System) in "title"					
Comment Web of Science Only relevant journals selected, limited to full text access and peer reviewed scholarly articles, only English and German, search period 01.01.2008 - 31.12.2018. keyword search in "Topic" only					
Comment ScienceDirect Only relevant journals selected, limited to full text access and peer reviewed scholarly articles, only English and German, search period 01.01.2008 - 31.12.2018. keyword search in title and abstract					

Appendix C: Case Study Interview Guidelines

Case Study Interview Guideline – Data-based applications in manufacturing

Note: this guideline has been modified for the appendix by reducing space for notes.

Data-based Applications in Manufacturing

0 Introduction

- Introduction of interview partner(s)
- Presentation of the research context and discussion of the interview objectives
- Interview procedure (Structure and duration of interview)
- Procession of data (confidentiality, anonymity, permission for audio recording)
- Further questions?

1 Lean Status Quo

- What is Lean for your company?
- Do you operate according to lean principles?
- What are key objectives of lean for your company?
- What are key objectives of lean for your company?
- Do you use the following lean practices?

Lean Practice	Yes - Notes
Preventive Maintenance (PM)	
Quality Management (QM)	
Continuous Flow Production (CFL)	
Pull System/Kanban (PU)	
Quick Changeover Techniques (QCT)	
Lot Size Reduction (LSR)	
Value Stream Mapping (VSM)	
Continuous Improvement (CI)	
Cross-functional Teams (CFT)	
Self-directed Work Teams (SDT)	

2 Digitalization Strategy

- Does a digitalization strategy exist for production (including logistics)?
- Is the integrated consideration of lean and digitalization desired?
- Is digitization and lean integrated or considered separately in your company?
  - Integrated strategy
  - Integrated organizational responsibility
- How does your company ensure that lean and digitalization is considered holistically?

3 Data-based application

3.1 Motivation

- Definition of smart manufacturing
- Link to results of Lean2020 study:

Derived research objectives

- Objective 1: Identify and describe data-based applications
- Objective 2: Investigate the impact of data-based applications on lean principles
- Objective 3: Identify challenges and enablers for the use of data-based applications.

3.2 Definition DBA in this research and overview DBAs

- Presentation of four DBA criteria
- Presentation of identified DBAs of literature review

Case Study Interview Guideline – Data-based applications in manufacturing

3.3 Data-based applications in practice

Which data-based applications do you use in your company?

1: Not in use2: In the testing phase
3: In use

DBA	1	2	3	Objectives and current status
Advanced Production Scheduling				
Layout Planning				
Real-time Control				
System Performance Measurement				
Predictive Maintenance				
Prescriptive Maintenance				
Material Flow Management				
Inventory Management				
Product Quality Monitoring				
Product Quality Improvement				

Further data-based applications currently in use

4 Company DBA use case

Please describe one or two application examples in more detail.

- What were or are the objectives of the application?
- What are the drivers for the application?
- What were or are obstacles and challenges?
- What are prerequisites and enablers?
- What are the lessons learned?

Notes:

5 Challenges and enablers for DBAs

Regardless of the specific application discussed above, what are or have been

- key barriers and challenges for the use of DBAs
- key capabilities and enablers for the use of DBAs

Notes:

6 Support potentials of DBAs for lean practices

- Underlying assumption: DBAs can support the effectiveness of lean practices
- Description of DBA – Lean Practice Impact Matrix and evaluation procedure
- Selection of 2-3 DBA – lean practice combinations with theoretically large impact or ambiguous impact (selected combinations depending on interviewee background).

Support potential Case 1:
Support of Lean practice *A* through data-based application *X*.

- Short description of DBA:
- Expected impact and theoretical reasoning:
- Feedback / evaluation of expected impact by interview partner (theory perspective and experience from the plant)

7 End of interview

- Miscellaneous / Administration / Confidentiality
- Interview Summary

Appendix D: Academic Expert Interview Guidelines

Academic Expert Interview Guideline – Data-based applications in manufacturing

Note: this guideline has been modified for the appendix by reducing space for notes.

Data-based Applications in Manufacturing

0 Introduction

- Introduction of interview partner(s)
- Presentation of the research context and discussion of the interview objectives
- Interview procedure (Structure and duration of interview)
- Procession of data (confidentiality, anonymity, permission for audio recording)
- Further questions?

1 Motivation for data utilization in manufacturing

- What motivates manufacturing companies to invest in data-based applications?
- Who is the key driver of DBA projects? (Initiate and managing the project)
- [Prof. Wuest only] According to the media, the USA and China are the leading nations in artificial intelligence/machine learning. What are differences between these countries and Europe regarding AI?

2 Data-based application

2.1 Motivation

- Definition of smart manufacturing
- Link to results of Lean2020 study:

Derived research objectives

- Objective 1: Identify and describe data-based applications
- Objective 2: Investigate the impact of data-based applications on lean principles
- Objective 3: Identify challenges and enablers for the use of data-based applications.

2.2 Definition DBA in this research and overview DBAs

- Presentation of four DBA criteria
- Presentation of identified DBAs of literature review

2.3 DBA Evaluation of DBAs

- based on personal experience in joint research projects with industry

DBA	Evaluation: E.g. - Current use in industry - Motivation and potential - Challenges and enablers in DBAs
Advanced Production Scheduling	
Layout Planning	
Real-time Control	
System Performance Measurement	
Predictive Maintenance	
Prescriptive Maintenance	
Material Flow Management	
Inventory Management	
Product Quality Monitoring	
Product Quality Improvement	

- Are there more DBAs you have worked on in your industry projects?

Academic Expert Interview Guideline – Data-based applications in manufacturing

3 Challenges and enablers for DBAs

From your experience in research and industry projects - what are or have been - key barriers and challenges for the use of DBAs - key capabilities and enablers for the use of DBAs
Notes:

4 Discussion DBA value model

- Presentation of initial DBA value model
- Presentation and discussion of underlying assumptions
  - Three levels (visibility, insight, action)
  - Value increases from level 1-3, significant value creation only on level 3
  - Effort comparably low on level 1, very high on level 2 and in between on level 3

Feedback:

Presentation and discussion of implications

- The situation of stage 2 is challenging for companies due to
  - Effort >> value
  - Duration of stage 2 not determinable due to uncertainty of data analytic results
  - Comparably high risk of "failure" of analytic DBA projects
- Managers strive to minimize risk → structural disadvantage for data analytic DBAs projects in competition for project funds.

5 End of interview

- Miscellaneous / Administration / Confidentiality
- Interview Summary

Curriculum Vitae

Name	Paul Johannes Bueß
Date & place of birth	December 18, 1988 in Sindelfingen, Germany

Work Experience

Since 11/2019	**Dürr AG** *Bietigheim-Bissingen, Germany* Consultant in the team Digitalization
11/2015 – 08/2019	**University of St.Gallen – Institute of Technology Management**, *St. Gallen, Switzerland* Research Associate in the division of Production Management
09/2014 – 10/2015	**KIT – Institute for Production Management (wbk)** *Karlsruhe, Germany* Student assistant in the area of quality assurance in the battery assembly
08/2013 – 12/2014	**KIC InnoEnergy SE** *Karlsruhe, Germany* Working student to COO
03/2013 – 06/2013	**EnBW Trading GmbH** *Karlsruhe, Germany* Intern in the department Analysis and Evaluation Fuels & Carbon
03/2012 – 08/2012	**Robert Bosch GmbH** *Karlsruhe, Germany* Intern in the department Customer Logistics and Planning
10/2010 – 04/2012	**KIT - Institute for Economic Theory and Statistics** *Karlsruhe, Germany* Tutor for lecture: Volkswirtschaftslehre I
05/2009 – 06/2009	**Daimler AG** *Sindelfingen, Germany* Internship

Education

07/2016 – 02/2020	**University of St.Gallen**, *St. Gallen, Switzerland* **Doctor of Philosophy in Management** focusing on Business Innovation
04/2015 – 10/2015	**Karlsruhe Institute of Technology**, *Karlsruhe, Germany* **Master of Science** in Industrial Engineering and Management
05/2013 – 10/2015	**Aalto University**, *Espoo, Finland* **Master of Science** in Industrial Engineering and Management
10/2012 – 02/2013	**Karlsruhe Institute of Technology**, *Karlsruhe, Germany* **Bachelor of Science** in Industrial Engineering and Management
10/2009 – 04/2013	**Johannes-Kepler-Gymnasium,** *Weil der Stadt, Germany* **Abitur** (German equivalent for A-level)