THESIS FOR THE DEGREE OF LICENTIATE OF ENGINEERING

Increasing the Value of Data in Production Systems

MAJA BÄRRING

Department of Industrial and Materials Science

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2019

Increasing the Value of Data in Production Systems
MAJA BÄRRING

Licentiate thesis at Chalmers University of Technology
Report no. IMS-2019-2

Department of Industrial and Materials Science
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover illustration:
Future technology with digital virtual connection
Purchased from Storyblocks

Printed by
Chalmers Reproservice
Gothenburg, Sweden 2019

ABSTRACT

A digital transformation is taking place, where information is available about almost anything, changing how work is performed and anticipated. More digitized information enabled by the digital technologies is supporting businesses to measure more about their processes and thereby also to know more. The same transformation is taking place in the manufacturing domain, which is referred to as the fourth industrial revolution. There are numerous national initiatives to approach the fourth industrial revolution and the aim is to make the manufacturing industry more digitalized and increase competitiveness. Digitalization is making more information about processes available, but it is first when data is informing decisions in an organization it will add value.

Along with all the benefits and potential values of the digital transformation, much of the attention has been on the technologies and systems that can enable the digitalization. Less focus has been spent on how the technologies and systems should be put into practice in the organizations to fulfill the needs of the manufacturing domain. New knowledge about digital technologies in combination with already existing expertise about manufacturing processes is needed. The aim of the thesis is to identify the value of data for decision-making. The approach outlined in this thesis will identify the values gained from the raw data itself and from the further processed data to provide decision support. The distinction between these two forms of data, raw data and further processed data, is important because it is believed that these can provide different values and that they involve different challenges for the organization.

5G telecommunication and 3D laser scanning serve as digital technologies in this thesis to enable more data in digital form on a production system. 5G was used for connecting a machine enabling the collection of data about critical machine components. 3D laser scanning was used to collect the spatial data in a factory environment. The results show that more data available about the connected machine provide values to the organization to know the status of the machine, be able to compare the designed system against the behavior in the real-world setting, a better understanding of the process and to learn from data. Spatial data provide values by being able to represent the production system as-is in a very accurate and photorealistic way. The values identified from having more data available for the decision support were in the daily operations to know the condition of the machine, for the manufacturing organization to plan proactive actions, and for the production engineer to understand the behavior of the designed system in the real-world context. The spatial data could both support when making changes to the physical setup and when planning the design of the factory environment in an offline mode.

The initial studies presented in this thesis supported to build the understanding of the current practice of data as decision support in the production organization. The understanding that data should support decision-making was high, but the data availability in the current state was scarce or of poor quality. This strengthens the aim of the thesis, to provide results that can show the value of data for decision-making. *"To measure more is to know more"* (McAffee and Brynjolfsson, 2012) is a statement serving as a cornerstone throughout this thesis and has also been justified by the results presented to answer research question 1 and 2. Data enabled by digital technologies can support multiple roles in the manufacturing organization throughout the different phases of the production system, for example in daily operations and maintenance.

Keywords: Digital technologies, digitalization, data, decision support, manufacturing, Industry 4.0, Smart Manufacturing

ACKNOWLEDGMENTS

This licentiate thesis compiles the work that I have done during the last 2,5 years as a PhD Student at the division of Production Systems at Chalmers. It is a milestone in terms of time since I am halfway through my PhD studies, but it has also been a time to reflect. As I decided to leave my role in industry to join Chalmers, I did not know much what I could expect. At that point, it was my interest in the production area that drove me to seek new challenges. Since then I have for sure learned more about the production area, but it has also a journey where I am starting to learn how to do research. This journey has been accompanied by numerous people and I would like to acknowledge those persons here for their contribution.

First and foremost, Professor Björn Johansson and Professor Johan Stahre. You were willing to give me this opportunity to join your group as I applied for the position in spring 2016. Since then you have challenged me both in my work and in my research. You have provided me with interesting opportunities and given me responsibility where you have trusted me to be able to follow through and succeed. At the same time, you have always been available as support and persons I could go to for advice and recommendations. Thank you!

My research and results have been dependent on the research projects that I have been involved in. The research project “The Space Industry of Tomorrow” in collaboration with RUAG Space AB and funded by NRFP have provided me with a lot of data and input to my research but also an understanding for the challenges in the space industry. Thank you, everyone, in the project team that has given tremendous support during the project. 5G Enabled Manufacturing (5GEM) in collaboration with Ericsson AB and SKF AB, funded by VINNOVA, has been a second research project that has been important for this thesis to happen. The learnings and outcome of this project have been really valuable for my own research but also to understand the potential of 5G for the manufacturing domain. Thank you, everyone, involved in the project and a special thanks to involved in the publications that are appended here. Currently, I am also involved in two new projects that will be important for my continuing research; the EU funded project BOOST 4.0 and VINNOVA funded project Digitala stambanan. Looking forward to continuing my work together with everyone involved in these two projects.

To learn more about the research process and how to pursue research, I would like to thank the graduate school Human-Technology-Design. The courses have definitely advanced my understanding of the research process and it has been a pleasure to do the course together with other PhD students at IMS. Another graduate school of importance for me has been Produktion2030. Besides the fruitful and interesting sessions that we have had, it has also broadened my research network in Sweden. A special thanks to the course P02 International Production, where we did an amazing trip together to Japan. It was hard work for the study trip to happen, but it was all worth it and it was a pleasure to do it together with all the participants.

Chalmers has a friendly and open atmosphere, which make me really enjoy to have it as my workplace. All my colleagues at PS contribute to this atmosphere as well as everyone else at IMS. Thank you! Besides my present collogues, I also would like to acknowledge all the people that I have got to know at my previous workplaces. I would like to thank both the collogues from my time at Siemens, both in Finspång and Charlotte, NC, and the year at the Student Union at LTH. I have learned so much from you.

As a last remark, I would like to thank all my friends that I have got to know at different points in life. You are sources for energy to me. Besides friends, my family is a tremendous support to me and has always been. You support me in whatever I do and push me to go even further. I am happy to be back in Gothenburg so I can spend more time with you again. Thank you!

Maja Bärring
Gothenburg, January 2019

APPENDED PUBLICATIONS

Publication 1:	**Bärring, M.**, Nåfors, D., Henriksen, D., Olsson, D., Johansson, B., & Larsson, U. (2017, 3-6 Dec. 2017). *A VSM Approach to Support Data Collection for a Simulation Model.* Paper presented at the 2017 Winter Simulation Conference (WSC).
Contribution:	First author of the paper, project member in the research project and supervisor for the master thesis that performed the study.
Publication 2:	**Bärring, M.**, Lundgren, C., Åkerman, M., Johansson, B., Stahre, J., Engström, U., & Friis, M. (2018). *5G Enabled Manufacturing Evaluation for Data-Driven Decision-Making*. Procedia CIRP, 72, 266-271.
Contribution:	First author of paper as well as developing and performing the study. Project member of the testbed project 5G Enabled Manufacturing (5GEM).
Publication 3:	**Bärring, M.**, Berglund, J., Johansson, B., Stahre, J., Iupikov, O., Alayon Glazunov, A., Ivashina, M., Engström, U., Harrysson, F. & Friis, M. (2019). *Digital Twin for Factory Radio Space Design of a 5G Network*. Submitted to Journal.
Contribution:	First author of the paper and one of the project members of enabling this demonstration of the testbed project (project B). Facilitating and providing the 3D environment to the study and leading the work of this demonstrator.
Publication 4:	Nåfors, D., **Bärring, M.**, Estienne, M., Johansson, B., & Wahlström, M. (2018). *Supporting Discrete Event Simulation with 3D Laser Scanning and Value Stream Mapping: Benefits and Drawbacks*. Procedia CIRP, 72, 1536-1541.
Contribution:	Member of the research project and acted as a supervisor during the simulation project and master thesis that collected input data to the simulation model.
Publication 5:	Åkerman, M., Lundgren, C., **Bärring, M.**, Stahre, J., Folkesson, M., Berggren, V., Engström, U., & Friis, M. (2018). *Challenges Building a Data Value Chain to Enable Data-Driven Decisions: A Predictive Maintenance Case in 5G-Enabled Manufacturing.* Paper presented at the Procedia 28th International Conference in Flexible Automation and Intelligent Manufacturing (FAIM), Columbus, OH, USA.
Contribution:	One of three main authors to the paper, mainly written parts to the introduction and provided insights for improvement and development of the paper. Project member of the testbed project which constitutes the case for this study.
Publication 6:	**Bärring, M.**, Flores-García, E., Johansson, B., Bruch, J. & Wahlström, M. (2018). *Challenges of Data Acquisition for Simulation Models of Production Systems in Need of Standards*. Procedia of 2018 Winter Simulation Conference (WSC).
Contribution:	First author of the paper and provided text to separate sections as well as ensuring the quality of the entire paper. To the study, contribute with one of the case companies that is part of an existing research project.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

3D	Three dimensions
5GEM	5G Enabled Manufacturing
CPS	Cyber-Physical Systems
CPPS	Cyber-Physical Production Systems
CSI-lab	Chalmers Smart Industry Laboratory
DES	Discrete Event Simulation
ERP	Enterprise Resource Planning
HMI	Human Machine Interface
IVSM	Improved Value Stream Mapping
IIoT	Industrial Internet of Things
ICT	Information and Communication Technologies
IoT	Internet of Things
IT	Information Technology
JSON	JavaScript Object Notation
LiDAR	Light Detection and Ranging
MOST	Mobile Operator Support System
MQTT	Message Queuing Telemetry Transport
noSQL	not only Structured Query Language
OPC UA	Open Platform Communications Unified Architecture
PLC	Programmable Logical Control
RCA	Root Cause Analysis
RPI	Raspberry Pi
TCP	Transmission Control Protocol
VNM	Value Network Mapping
VSM	Value Stream Mapping

LIST OF FIGURES

LIST OF TABLES

1

INTRODUCTION

The introduction will initially give a background to the research area followed by stating the research gap and problem formulation for this thesis. The vision and aim for the research will be introduced leading into the defining the research questions. Further on will stakeholders to this thesis be stated, the delimitations clarified, and the structure of the remaining of the thesis explained.

1.1 BACKGROUND

The digital transformation taking place is changing how work is performed and anticipated in almost every industry (“5G systems”, 2017). Everything which can gain value from having a network connection is expected to have one in the future (“More than 50 billion devices”, 2011). The almost daily advances in technology and techniques that mean that more details of companies’ operations can be managed (Davenport, 2006). More digitized business activities will entail new sources of information and, aided by ever-cheaper equipment, will bring society into a new era. This new era will comprise a large amount of digital information available on various topics of interest for the business. With more digitized information and sources, more can be measured about processes and more about the business can therefore also be known (McAfee and Brynjolfsson, 2012). However, not only is more information available than ever before, but information is also growing at a faster pace than has ever been experienced and the change in quantitative measures is becoming a qualitative change (Mayer-Schönberger and Cukier, 2013). The drivers for businesses to connect their devices will be a customer-service enhancement and the potential of a more cost-efficient and productive organization, as well as the ability to control and manage assets (“More Than 50 Billion Connected Devices”, 2011, Ericsson White Paper). In traditional sense, possessing knowledge meant that the past could be understood. Now, however, with all of the data available on almost anything, possessing knowledge is more about being able to predict the future (Mayer-Schönberger and Cukier, 2013).

Digitalization is bringing profound changes to businesses, in every industry, regarding how operations and work can be performed. The manufacturing domain is no exception and the fourth industrial revolution, or alternatively the second machine age (Brynjolfsson and McAfee, 2014), has been coined as an expression to describe this. The national initiatives to address this topic are many, such as Industrie 4.0 in Germany, the Industrial Internet of Things and Smart Manufacturing in USA (Kang et al., 2016). The aim is to address the changes required in order to make the manufacturing industry more digitalized, and hence more competitive. All of these initiatives will combine digital connectivity, advanced analytics, artificial intelligence, and new cyber-physical systems for the realization of the physical goods as the output of the factory (Behrendt et al., 2017). Installed sensors in the manufacturing processes will provide a constant stream of data (Bughin et al., 2015). Communication between humans, machines, parts, and products will become close to real-time and data will be stored in the cloud, increasing the availability and accuracy of it. Both expected and unexpected changes in the organization will be handled as a result of more flexibility in the organization (Rüßmann et al., 2015). As the enabled data can inform decisions in the organization, this is when real value and competitive advantages can be gained by the organizations that implement and integrate the new technologies (Bughin et al., 2015).

1.2 RESEARCH GAP AND PROBLEM FORMULATION

The amount of internal and external data available today in organizations has increased as a result of digitalization. The urge in the organization to make something useful of the data has therefore also become more important (Gillon et al. 2012) and organizations are starting to view data as an asset (Khatri and Brown, 2010). So far much of the attention has been on the technologies and systems that enable digitalization and less on how these should be used in practice in the organizations (Gillon et al., 2012). Applying the big data concept has potential benefits in the decision-making process but it will not guide the decision itself (Poleto et al., 2015) and it can only be leveraged when it starts to drive decisionmaking (Gandomi and Haider, 2015). Data itself does not necessarily create value for the organization, but rather the more informed decision that can be made with more data to support it (Chatfield et al. 2015; Li et al. 2014). As an organization is transforming to a more data-driven state, it becomes more important to have high-quality data available that can be trusted by the organization and that describes the reality of what is attempted to be described by data (Brous et al., 2017). To transform to this state, which includes the design and building of the data value chain, requires new knowledge about digital technologies in combination with the already existing expertise about the manufacturing processes. There is a need to show how data provides value to the organization as decision support.

1.3 VISION FOR THE RESEARCH

The vision for the research is to show to the manufacturing field the applicability of the digital technologies that can digitalize data describing the production system and how benefits can be gained by supporting decisions with data. The future manufacturing company will need to be driven by fact-based decision support based on data instead of relying solely on intuition and experience. In order to achieve this vision, the research will provide studies that can prove the capability of digital technologies to provide more data and evaluate how making more data available brings values to the manufacturing organizations.

1.4 AIM OF THE LICENTIATE THESIS

The aim of the thesis is to identify the value of data for decision-making. The approach outlined in this thesis will identify the values gained from the raw data itself and from when this data has been further processed to provide decision support. The distinction between these two forms of data, raw data and further processed data, is important because it is believed that these can provide different values and that they involve different challenges for the organization

1.5 RESEARCH QUESTIONS

This work relies on the precondition that data in a digital form that describes the reality of the production system has been scarce until the present day, which has hampered organizations in making informed decisions based on data. To state the current practice for data, the thesis will exemplify the availability of data about the production system, challenges of data acquisition and the understanding by manufacturing personnel of the future use of data. To explore the values that can be provided by having more data available for decision-making, the current situation and the future situation, with more production data, will be explained for each study.

Digital technologies are used to capture and digitalize data from the workshop. Data describing the reality can be quantifiable facts about an event that has occurred, or spatial data describing the geometry of the workshop. Data can also effortlessly be copied, combined with other data, and can be used by multiple persons at the same time (Skoogh et al., 2012; Manyika et al., 2011). More exact data becoming more readily available should give insight to the organization for decision support. Organizations should be able to rely less on intuition and more on data. Figure 1 illustrates the process with data as the core.

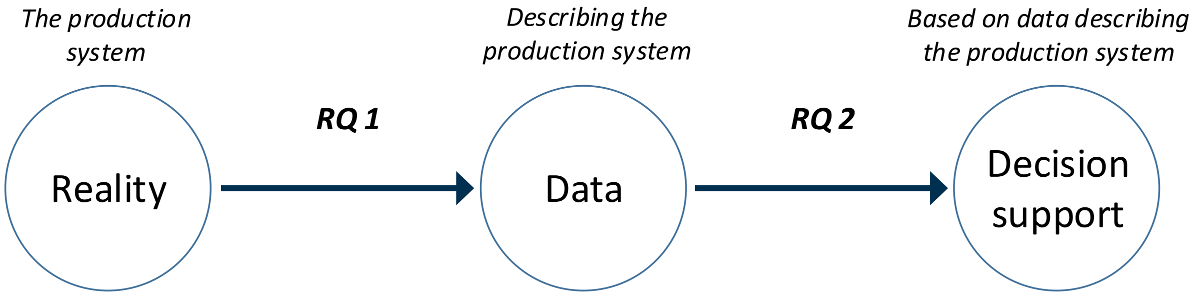

Figure 1. Summary of the research scope of the thesis

This thesis will focus on digital technologies that are able to provide an abundance of data to make information about the reality of the production system available and to support decision-making. The two research questions that will be addressed in this thesis cover firstly the data made available by digital technologies to describe the reality, and secondly the transformation of data to provide decision support that will guide the human when making decisions. As Figure 1 illustrates, both RQ1 and RQ2 focus on the move from one state to the next. Since it is believed that the organization can gain value from both steps of the data process, this will be explored in each respective research question.

RQ 1) What is the value of collecting data from a production system using digital technologies?

RQ 2) What is the value of decision support based on an increasing amount of digital data from a production system?

1.6 STAKEHOLDERS

As was explained in the background section, the digital transformation is impacting on multiple instances in society and has the potential to change how things are done. This holds true both within the industrial domain and in other areas. The thesis will present recent literature references from both the research community of manufacturing and from areas where the advancements in technology and application of big data are discussed. This is done to advance the manufacturing research, so the main approach of the thesis is the research community within manufacturing.

Another stakeholder group that this thesis is intended for are the practitioners in the manufacturing sector. The intended purpose is to show and identify values that can be gained for the production organizations as they start using more data as decision support. The real-world cases presented in the thesis and in the appended papers serve to present results gained from the research that can support the applicability and value of having more data available. The cases studied also support the understanding of how digital technologies can be used for making more data available.

1.7 DELIMITATIONS

The frame of reference will spend a whole section on the topic of big data, and this will also be a recurring topic throughout the chapter. One reason why this is given a lot of attention is the vast literature that now exists on the topic in comparison to traditional data analytics. Although the big data and technologies explained in the frame of reference build the cornerstones of this thesis, it should be stated this is not a delimitation. Small data or any type of data that can provide more informed decisions are of relevance. It should also be stressed that the focus is not to explain how data analytics are performed in detail, and machine- and deep-learning algorithms are excluded from the content.

This thesis aims to provide results that can strengthen the idea that a more informed decision can support human decision-making in the production system. The intended recipient will be a human who has an interaction with the production system and needs the information for making a required action or change to the system. There will, therefore, be no results describing how more data can support the automated decision-making of machines. The thesis will also stop at the instance where data has become decision support since decision-making is a massive research area by itself and has been researched by numerous researchers before. The feedback loop to reality, when a change has been introduced to the production system to understand its impact, will also be excluded from the scope of the thesis for the same reason.

1.8 STRUCTURE OF THE THESIS

The structure of the thesis is to first present the theoretical framework that the thesis relies upon. This is presented in Chapter 2. The chapter covers the areas that are included in Figure 1, namely the reality, i.e., the production system, the data that is enabled and what the data sources in a production system are, the information that constitutes the decision support, and also the digital technologies that are the enablers in this thesis. Chapter 3 thoroughly explains the research approach that has been applied in the thesis and the individual studies in order to answer the research questions asked by this thesis. Chapter 4 explains the results from the studies that have been performed, divided into the different areas that have been studied. Chapter 5 answers the research question and in Chapter 6 the theory and results are combined in the discussion. Chapter 7 ends the thesis with the conclusions.

2

FRAME OF REFERENCE

Frame of reference is presenting related work and references in the field that the thesis relies on. To motivate the importance of making use of data, the first section will present related work in the area of data-driven decision-making. Further on will the terminology be presented to give a background and understanding of what can expected form the remaining of the chapter. According to Figure 2, moving from left to right, will the chapter present the reality that in this thesis is a production system. Data and information will be presented including the concept of Big Data and data sources in a production system. Thereafter will the digital technologies be explained going through the different areas that are in the lower level of Figure 2. Lastly, decision support will be defined.

2.1 DATA-DRIVEN DECISION-MAKING – RELATED WORK

"You can't manage what you don't measure" begins the well-cited paper *Big data: the management revolution* (McAfee and Brynjolfsson, 2012), where the authors let the meaning state the potential of the big data era for decision-making. "To measure is to know" is another saying to underline the same idea, that a phenomenon that can be measured can be understood (Mayer-Schönberger and Cukier, 2013). Previously data has been costly or difficult to extract, which has motivated basing decision-making on the intuition of people, meaning that they state their opinions about the future and act according to this intuition. Current possibilities for making use of data call for a new way to make decisions. The two MIT scholars McAfee and Brynjolfsson, authors of the paper *Big data: the management revolution*, describe the idea in brief as "data-driven decisions are better decisions – it's as simple as that." As data becomes available to managers, they can make decisions based on the evidence of what is happening in reality rather than on some peoples' experience and intuition. The paper takes a positive stance towards the benefits that more data can have for organizations in decision-making. The paper has been well cited within the research community and provides real-world cases that can strengthen this idea. It relies, however, on the precondition that technology development has provided us with more information available in digital format.

The World's Technological Capacity to Store, Communicate, and Compute Information, written by Hilbert and López (2011) published in Science magazine, has made an estimation of the technological capacity to store, communicate, and compute information in the world. This has been done by tracking these abilities for 60 analog and digital technologies over the time period spanning 1986 to 2007. In this context, storage is defined as the maintenance of information, communication as the amount of information that can be sent and received over a considerable distance, and computation as the meaningful transformation of information. During this period for the study, the capacity of the general-purpose computer was growing at an annual rate of 58%, which implied that in 2007 it was possible to store 2.9 × 1020 optimally compressed bytes, communicate almost 2 × 1021 bytes, and carry out 6.4 × 1018 instructions per second on general-purpose computers. At the same time, the annual growth rate for the capacity of bi-directional telecommunication was 28% and the globally stored information was increased by 28% per year.

McAfee and Brynjolfsson (2012) have a very positive stance towards the “data-driven decision-making” phenomena and Hilbert and López (2011) states the technological development increasing the world’s capacity to handle more digital information. The paper *Big data: a fashionable topic with(out) sustainable relevance for research and practice?* (Buhl et al., 2013) takes a more neutral stance in discussing what the potential benefits of big data are or if this is only a new buzzword for businesses. The term big data is a well-recognized term nowadays, both within academia and among practitioners, and of the publications dealing with the big data concept in 2012, 70% were published in the last two years (Pospiech and Felden, 2012). There is excitement about what big data can bring to a diverse set of areas, such as within science, governments, and industries. Despite the expressed excitement within all of these fields, a question raised by the authors is whether big data really is bringing something new or is just a new word for something data analytics have been doing for decades already? What has been seen from the analyses performed with big data is that it brings new challenges regarding volume, velocity, variety, and veracity of data that have not yet been solved. The authors also refer to both previously mentioned papers (McAfee and Brynjolfsson, 2012; Hilbert and López, 2011) to emphasize the potential of the big data concept. So instead of being viewed as a threat, the authors conclude that the big data concept should be handled as an opportunity for both academic and practitioners in further development.

2.2 TERMINOLOGY TO BE USED

Figure 1, presented in the introduction, contains the core of the thesis with the reality, data, and decision support. The reality in this context is defined as the production system, data as the digital representation of the production system that can describe its behavior and condition, and decision support is what should be based on the more digital data made available. To guide the content that will be presented in this chapter, Figure 2 is presented and explained. This figure visualizes the reality, data, and decision support and also the activities that take place in-between the three stages necessary for the realization.

The different steps of enabling more data, in the lower level of Figure 2, are adapted from literature about handling big data. This selection is made to support explaining what happens in these steps and how the technologies support in handling data, but this thesis is not limited to big data since any data that can support decision-making is of interest. Data sources and generation are data of primary or secondary type and new sources that can be extracted by technologies implemented in the production system, for example by sensors. Collection of data is where the generated data is made available digitally and transmitted, i.e., the communication of data by the digital technologies. Communication is the digital technology that supports from the stage where data is being collected to the end stage where the data is consumed by recipients. Storage of data is where data is maintained when not used and enables digital data to be available. For transformation from data to decision support, a computation, i.e., pre-processing and analysis takes place and this is where the data is transformed into insights and information, which are presented by a carrier and as content to a receiver of the information. What has been explained in brief in this paragraph will be explained more comprehensively in the remainder of this chapter.

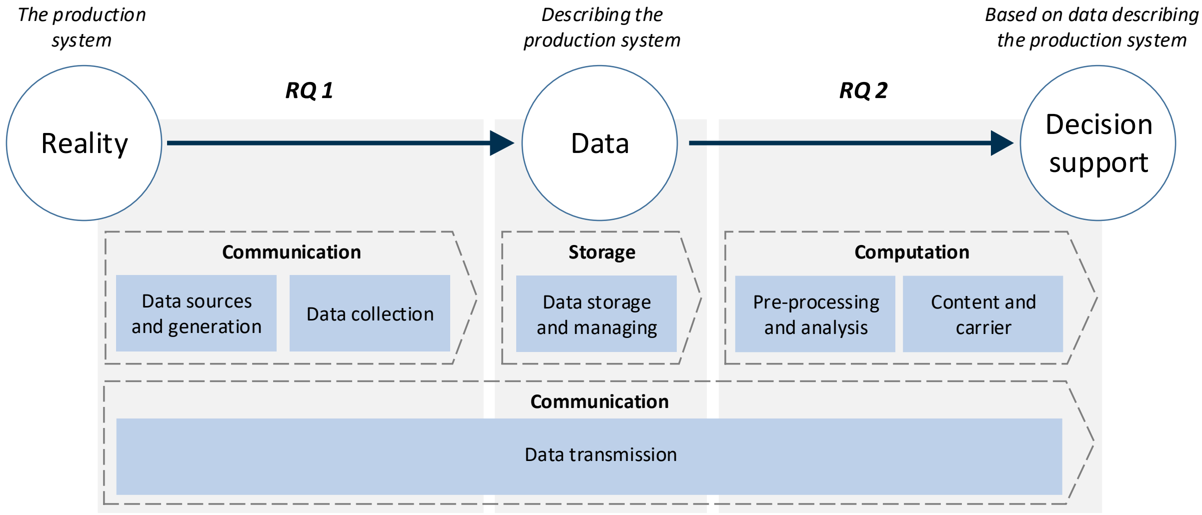

Figure 2. Summarizing and connecting the concepts that are used in the thesis

2.3REALITY - THE PRODUCTION SYSTEM

In order to define the production system, systems theory has to be introduced. Systems can be identified everywhere and even though most systems differ from each other there are some key components that unite them and make them possible to describe with the same terminology regardless of scientific area (Skyttner, 2005). A system consists of separate elements that interact with each other and together perform an intended function (Kauffman, 1980). The elements are designed to fulfill the functions based on the desired objectives and purpose of the system (Frezzini et al., 2011). The system boundary can be set at different levels and it is first when the system boundary is identified that it can be regarded as a system. Everything outside the system is viewed as the external surrounding which can impact on the system's goal but the system itself cannot impact on its surrounding (Churchman, 1968). The purpose of a system is defined by its function, which for the production system is the transformation of material to a product or service (Hubka and Eder, 1988; Bellgran and Säfsten, 2005).

Production is the transformation process where an input that goes into the system is transformed to an output (Wu, 1994). It is a process of combining material, resources, labor and capital in order to create products and/or services (CIRP, 1990; Jonsson and Mattsson, 2009). A number of areas are required for the transformation; technology, humans, energy, and information need to be organized and managed in an effective way to make the transformation possible (Bellgran and Säfsten, 2005). It is a sequence of value-adding activities and operations where the material is transformed to reach the final desired state (Jonsson and Mattsson, 2009). The production system requires a holistic perspective (Rampersad, 1994; Wu, 1994; Bellgran, 1998) and the sub-parts of the system contribute with their internal relations in order to realize the transformation (Bellgran and Säfsten, 2005). As the system is designed with a holistic perspective in mind, consideration is taken to all the sub-parts of the production system that implies both the technical and physical aspects as the human in the system and how work should be organized (Bennett, 1986). Facilities, humans, and equipment (e.g., machines), software and procedures are considered as elements of the production system which all have relations to each other (Löfgren, 1983; Chapanis, 1996). The transformation process in the production system is visualized in Figure 3.

Both manufacturing and production systems are used terms to refer to the transformation of products, but there is a distinction between them. Manufacturing production, or simply production, is the physical acts and processes where the material is transformed into a product. The definition of manufacturing, on the other hand, does not only include the physical transformation of material to a product, but also the interrelated activities that need to be performed in order to realize a product and get it out to the market. These activities are the design, selection, planning, production, quality assurance, management and marketing of products. In that sense, the manufacturing term can be viewed as overarching the definition of production and involves all activities within a company that realizes the product

(Hounshell, 1984; CIRP, 1990).

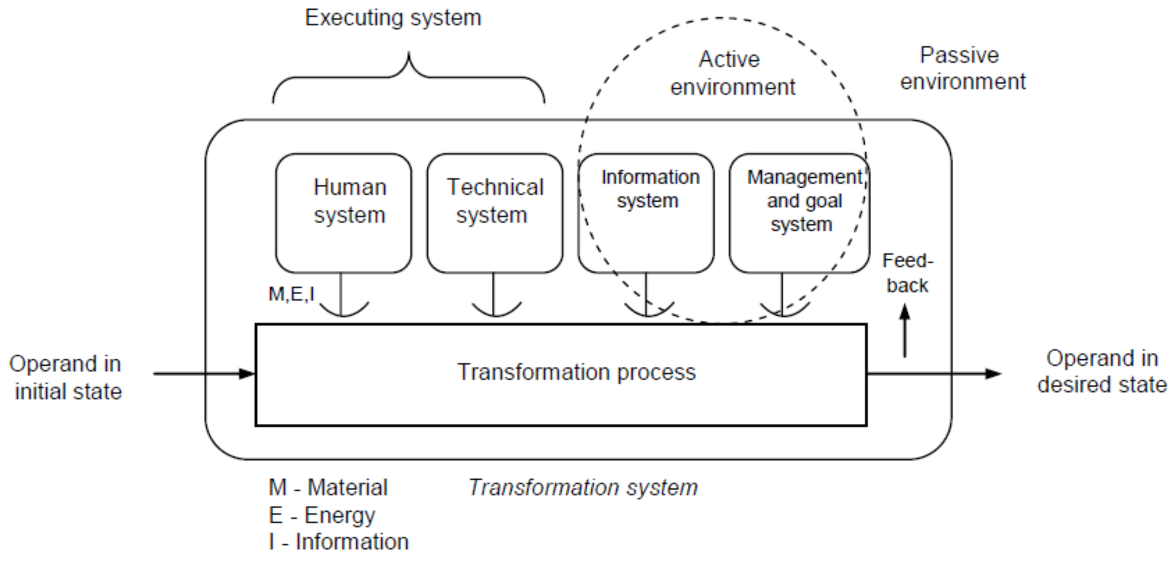

Figure 3. A simplified model of the transformation system (Hubka and Eder, 1988)

Until now, there have been three industrial revolutions; mechanization through water and steam, mass production in assembly lines, and automation using information technology. A fourth industrial revolution is now expected and it is termed Industrie or Industry 4.0 in Germany and Smart Manufacturing or Industrial Internet of Things in the United States. The complexity has increased for every revolution in the industrial domain, which needs to be managed by the right methods and tools. These initiatives describe the transition toward a heavy data focused, supply network-wide integration of information and communication technology (ICT), and increased automation with the human still involved in the system (Thoben et al., 2017).

Industry 4.0 describes the use of ICT in the industrial context and the idea is to make use of data to generate knowledge to decision support in all parts of the business (Schuh et al., 2017). Flexibility, speed, productivity and the quality of the production processes have the potential of being improved with Industry 4.0. The previously isolated, optimized cells will come together as a fully integrated, automated and optimized production flow as a result of the advancements in technology that Industry 4.0 rely on. This will change the relations between suppliers, producers, and customers as well as between the humans and machines in the production and will bring new values. The collection of data will be from various sources such as production equipment and systems as well as enterprise- and customer management systems. The extensive evaluation of the data will also be a decisive part to support decision-making in real-time (Rüßmann et al., 2015).

To optimize the use of resources for the realization of high-quality products as well as the ability to quickly adapt to changing conditions in the environment will be key for the realization of future manufacturing (Lu et al., 2015). The Smart Manufacturing Systems (SMS) involves the use of advanced technologies that can handle the rapid flow and widespread use of digital information that exists within a manufacturing system as well as between manufacturing systems (Shipp et al., 2012; Davis et al., 2012; SMLC, 2011). This is done with the aim to maximize the capabilities of a manufacturer that is critical for competitiveness, namely the cost, delivery, flexibility, and quality (Rodrique et al., 2014). SMS applies ICT together with software applications that can support in the optimization of labor, use of material and energy for the production of products with the attributes of high-quality, high customization and on-time delivery. A vision for SMS is that the product by itself will provide information telling the history of how, when, where it was manufactured (Lu et al., 2015).

Internet of Things (IoT) and Industrial Internet of Things (IIoT) rely on the standard technologies to connect and populate devices, or in some cases even unfinished products, with embedded computing capabilities. Field devices will thereby be able to communicate and interact with each other and features such as responses in real-time, decentralized analytics and decision-making will become possible to support (Rüßmann et al., 2015). The “smart” devices equipped with sensors and the ability to communicate seamlessly over the Internet with other devices or the cloud are preconditions for IIoT. Improved infrastructure for enhanced connectivity will also support IIoT. Two important features of IIoT will be (1) the benefits of the cloud that can provide robust storage and computing power at low cost and (2) that the devices (edge) can process and store data. Sensors implemented in the production line will inform the manufacturing personnel about the status of the machine (Patel et al., 2017).

Cyber-Physical Systems (CPS), or also called connected systems, are systems that can interact with each other over the standard Internet-based protocols. They can analyze data in order to be able to predict failures, configure themselves and perform required actions to adapt to changes (Rüßmann et al., 2015) and in the core of this concept lies the use of smart devices (Lu et al., 2015). CPS are systems of computational entities, which are in connection with their surroundings, i.e., the physical world and are both providing and using data-accessing and data-processing services (Monostori, 2014). For the manufacturing system, the concept is more specifically termed as the Cyber-Physical Production System (CPPS). The development of production modules that will incorporate the smart devices will be realized with a reference architecture for CPPS. This together with other efforts will allow transparency of data in the system from the lowest to the highest level in the manufacturing system (Lu et al., 2015). The CPPS will have a significant impact on how the communication is carried out and it will be situation dependent and across all levels of the production over-riding the current hierarchy that exists in the system (Monostori, 2014).

2.3.1 Values in the Production System

Adam and Swamidass (1989) present a review of the literature in the field of manufacturing strategy research. The content variables that were mentioned most frequently among the references, both when it comes to empirical and conceptual writing, are presented in Table 1, adapted from Adam and Swamdiass (1989). They are cost, flexibility, quality, and technology-process (manufacturing process and manufacturing technology are here seen as interchangeable and the term technology-process is here used).

Table 1. Adapted from Adam and Swamidass (1989) with the four most frequently mentioned content variables for manufacturing strategy

References (Authors and year)	**Content variables**			
	Cost/price/ productivity	**Flexibility**	**Quality**	**Technology-process**
Hill (1985)	X	X		X
Miller and Roth (1988)	X	X	X	
Schroeder et al. (1986)	X	X	X	X
Skinner (1978)	X	X	X	X
Swamidass (1986)	X	X	X	X
Wheelwright (1984)			X	X

Cost is a strategic variable (Porter, 1980; Skinner, 1978) and in the manufacturing context is productivity a common substitute for it. This is because it more accurately puts a value on the efficiency in the transformation from input to output (Adam and Swamidass, 1989). Productivity can be applied at multiple levels; at the wider national level to the small-scale level to quantify employee productivity and everything in between (Kendrick, 1986). Besides the productivity and cost, quality is also an important aspect for the profitability of a company which is the degree of excellence a company manufactures products and can be divided to process and product quality (Son and Park., 1987). The common view of quality has been the classical cost-quality trade-off. Cost related to the quality of products are of the type prevention cost and to achieve competitive advantages when it comes to quality is to work with the prevention costs (Evans & Lindsay, 1989). Flexibility tells to what degree a manufacturing company can adapt its production system to changes introduced and can involve equipment, product, process or demand flexibility to mention some (Son and Park, 1987). The last content variable relates to technology-processes and already in 1989, when Adam and Swamidass (1989) presented the review of literature in manufacturing strategy, the topic of advancements in technologies and their justification was discussed.

2.4 DATA AND INFORMATION

Data is a set of discrete, objective facts about an event that has occurred, which can be quantified, e.g., an event describing a cycle in a machine (Skoogh et al., 2012). To distinguish it from information, it can be explained as the raw material for the creation of information (Davenport and Prusak, 1998). Data is the raw indicator of an event, while information is the meaningful interpretation of the signals that provide insights for action. What differs data from other assets is its ability to effortlessly be copied, combined with other data and be used by multiple persons at the same time (Manyika et al., 2011). Data can still contain value even when it has been used and this is because it can be used over and over again. It does not wear out as it is being used by multiple persons, which is called "non-rivalrous" good. The value of a dataset may be gained first when it is combined together with another one, even a very different one. It is, however, the primary use of data that motivates and justifies the collection and processing of it, but if a secondary use is considered already at the collection stage this can support making the data suitable for multiple purposes, called extensible data. Even if the amount of value gained from its first use is low, the total value can be big when the data is exploited effectively (Mayer-Schönberger and Cuiker, 2013).

Information is a message (Davenport and Prusak 1998) from further processed data (Skoogh et al. 2012). By further processed, it implies that the data can be contextualized, categorized, calculated, and condensed. Information is data that has gained a purpose and relevance for the context and it is meant to impact on the receiver's understanding of something (Davenport and Prusak 1998). By gaining information the uncertainty about the condition of reality can be decreased and from an information theory perspective, information is defined as the opposite of uncertainty (Shannon, 1948).

Data and information are no new phenomena, but what has previously been constraints for the collection of data do not longer exist to the same degree and that is what makes the current era different. Analyzing data has according to traditional measures involved to use statistical methods to analyze the data available. This involved to concentrate, extract, and refine the useful data among all the data available in the chaotic dataset and to extract the value of the data (Chen et al., 2014). The technology available today has reached a point where most data can be captured and recorded to a cost that is justifying the same. And at the same time as the power to process the data is increasing, the amount of data available is doing the same (Mayer-Schönberger and Cuiker, 2013). In the next section, the big data term will be further explained and explored.

2.4.1 Big Data

With more operations and processes becoming digitalized, more data has also become available and is often referred to as big data. The urge to make use of big data has grown rapidly during the last years and the interest for it is present within both academia and industry. Since data has increased in terms of volume in such an explosive way lately, the term of big data has mainly been used to describe the enormous datasets that need to be handled. The new datasets in the big data era typically include masses of data that are unstructured in character and need to be analyzed in a timelier manner hence in real-time compared to the traditional data analysis. This situation is providing new opportunities for discovering new values and to gain an in-depth understanding of the previously hidden values in the data. The new massive datasets available also imply new challenges, for instance how to handle and organize the data in an effective manner. Technical speaking the big data term should be used for the datasets where the size of them cannot be perceived, acquired, managed or processed by the traditional IT and software/hardware within an acceptable time frame (Chen et al., 2014). It comprehends the collection, storage, transportation and exploitation of big data (Zhou et al., 2014).

Besides the volume and large data sets that are related to the big data concept, it is also about the variety and velocity of data. The Gartner report from 2001 authored by Doug Laney is well-cited and has defined big data based on three Vs of data; volume, velocity, and variety of data (Laney, 2001). Lately, other authors have added more Vs to the definition, such as value, variability, and veracity of data (Gandomi and Haider, 2015). Limiting the definition to the three Vs, volume implies the amount of data that is created which is something that has grown enormously and still is doubling every 40 months (McAfee and Brynjolfsson, 2012). The velocity of data incorporates that data should be retrieved fast, even in real-time for some cases, but also analyzed and acted upon in a fast manner (Gandomi and Haider, 2015). As the data become available in real-time, the organizations can respond to changes in a timelier manner and in some cases, the business success can be more dependent on how fast the data can be received than the volume of data (McAfee and Brynjolfsson, 2012). More sources of data are becoming available, enabling data that are both unstructured and heterogeneous. This is referred to as the variety of data and the potential will be to make use of data in the form of messages, updates, images, reading from sensors, the location from cell phones etc. (Gandomi and Haider, 2015; McAfee and Brynjolfsson, 2012).

In *Big data: A Survey* (Chen et al., 2014) the authors review the background and state-of-the-art for big data including the related technologies. The value chain of data explained in this paper, constitute four steps; the generation, acquisition, storage, and analysis of data. The data chain starts with the raw data material, that in the utilization process are generated and acquired, stored in the storage step and analyzed in the production step where the raw data are transformed to provide new values. The acquisition, in turn, involves three additional steps, namely the collection, transmission or transportation, and pre-processing of data. When acquiring big data, it involves new requirements for managing the data. As the data has been collected, a transmission mechanism needs to be utilized for efficiently sending the data to a storage management system that can support the different analytical applications for data analysis. In terms of capability of the storage system, the data center need to provide powerful backstage support since the big data concept has more stringent requirements on storage, processing, and network transmission capacity (Chen et al., 2014), see Figure 2.

Big data is, however, a resource and tool that is intended to be used for informing rather than explaining. Because depending on how well or poorly it is used, it can lead to an understanding or misunderstanding by the receiver of what the data tells. Big data is frankly a tool that can provide good-enough answers about the reality, but at the time being with the methods available, it is not giving any ultimate answers. What can be collected in terms of data will always only be a fraction of what actually exists and the real revolution of using more data is in the data in itself and how it is used. Humanity has long had the interest to quantify and understand the world. Big data makes an important step forward in this quest since things that never before has been possible to measure, store, analyze or share is now becoming datafied. The digital age, or information society, that was discussed in the introduction may finally reach its full potential with the big data era and encompass what the name is suggesting with the data in the center stage (Mayer-Schönberger and Cukier, 2013).

2.4.2 Data Generation – Sources of Data in the Production System

Data sources that are commonly used to describe the production system can be of either of primary or secondary type, see Figure 4. Measurements of discrete events in a production system recorded by the personnel for, e.g., a simulation purpose are termed as a primary source and the secondary data sources are the data that have been collected for other purposes. The secondary data sources include an external reference system, corporate business system, and project-specific data. Since data available in these sources have been collected for other purposes they can require some more processing before being used for the specific case (Skoogh et al., 2012) If data cannot be found in these sources for secondary data, data need to be collected, i.e., the primary data type (Skoogh et al., 2012).

For the simulation area, the absence of available data describing the process as well as the lack of quality in the data has been a known problem for long. Robinson and Bhatia (1995) have introduced a way of classifying data for simulation models based on the availability of data. The three categories are (Robinson 2004; Skoogh and Johansson 2008):

(A) *Available*: data is available, fits a simulation purpose, but requires validation.
(B) *Not available but collectable*: requires the investment of resources for its acquisition.
(C) *Not available and not collectable*: data should be based on an expert's opinion, similar production processes, or historical data.

Besides the availability of data, another important aspect for data to provide decision support with high quality is the challenges that are involved with data acquisition. They have been introduced by Robertson and Perera (2002) and modified by Skoogh et al. (2012). (1) Accuracy is to what degree data is free from errors and requires investigation of the source and format of data. If a lack of (2) correctness appears it is due to lack of standards, communication problems, or because of incorrect labeling of data. The (3) duplication of data implies that two or more sources hold data about the same event. As different sources of data can hold data about the same event, (4) consistency of the data can become a problem and thereby imply that the sources have different values for the same event. The (5) timeliness of data can post challenges for the data acquisition since it means that the data item will change value over time and therefore becomes invalid after a period of time. (6) Validity is a problem when the data is not describing the behavior of the real-world system and (7) reliability when the data is not trustworthy to the stakeholders. When data is not (8) complete, assumptions need to be made or additional data will have to be collected.

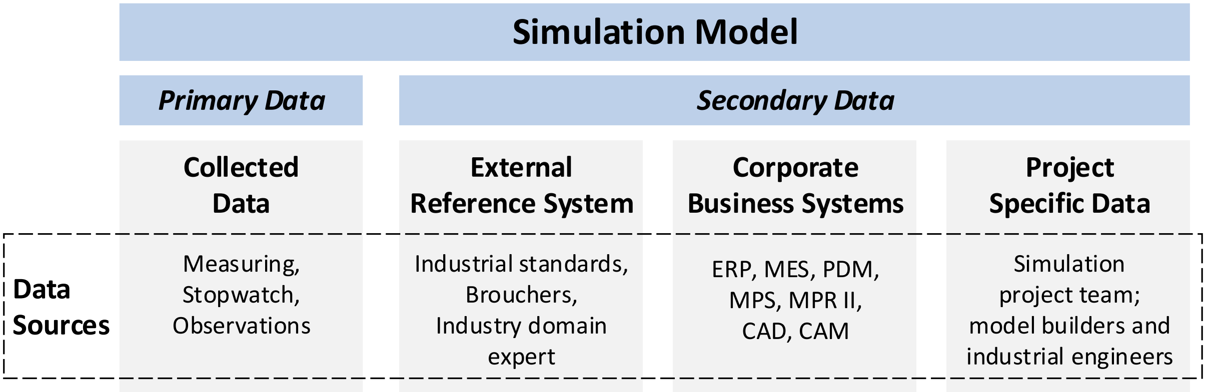

Figure 4. Categorizing data sources used for simulation projects (adapted from Skoogh et al. 2012)

Quality of data is a multidimensional concept impacting on the properties of information such as accuracy, timeliness, completeness, consistency, relevance and also fitness for use (Miller, 1996; Strong, Lee and Wang, 1997). Cleaning data is a process where inaccuracy, incompleteness or unreasonable data are identified so it can be modified or removed with the purpose to improve the quality of data. A common imperfection in datasets is the redundancy, repetition or surplus of data, which can increase the expenses when it comes to transmission of data and require large space for storage. This, in turn, can for the quality of data lead to inconsistency, reduction in data reliability and also damage the data (Chen et al., 2014).

According to the report from McKinsey (Manyika et al., 2011) the manufacturing sector, including both discrete and process manufacturing, have together the highest aggregated data stored in bytes. The intensity of data, however, ranks lower due to that they are divided by a large number of individual firms that do not share their data. To what extent they can manage to pool data across the manufacturing supply chains will be a decisive factor for if they will gain the potential value that lies in big data. Besides the volume of data that differs between the sectors, the type of data does as well. It can be data from video, images, audio or text/numeric and the manufacturing industry generates mainly from a historical perspective up until present, text and numerical data from their production processes. In addition to that is the R&D and engineering departments that are users of image data for design and development (Manyika et al., 2011).

2.5 DIGITAL TECHNOLOGIES

Digital technologies have had a tremendous impact on how society functions, how we live our daily life's and how businesses operate. This has stimulated economic growth and supported to increase productivity. Examples of digital technologies are computer, robots, cell phones, digital watches, traffic lights to mention some. The common denominator for them is the integrated computer chip that controls and runs them. A key characteristic of the computer chip, that has made it so successful, is that it can be integrated into a wide range of different devices in many different ways (Cortada, 2004). Computers brought devices for digital measuring and storage that made the datafying much more efficient and supported the mathematical analysis to uncover the unhidden value in the data (Mayer-Schönberger and Cuiker, 2013). As the computers were brought into the manufacturing domain in the second half of the twentieth century, this has had an enormous impact on how goods are manufactured and distributed. Computers have now become integral to all features of manufacturing (Cortada, 2004).

Besides the impact the computer chip has hade by itself, it has also stimulated the convergence of different technology areas that have created new levels of capabilities, functions and additional requirements. Examples of these parallel development tracks are telephony, telecommunication, radio, television, and computers. The digital technologies have become a base for the information technologies (IT) and devices, which has also enabled the possibility of merging them together. As has been seen in the development of digital technologies, the ability to store and handle a large amount of information and the speed of performing functions have improved at the same times as the cost for them have decreased. The smaller and less heavy technologies today is a change that is visible, but the reliability and ability to handle more data and complicated instructions have also made major leaps at the same time that is not as visible for the observer (Cortada, 2004). PCs are now available as smart devices and the transformation of the hardware has taken place in parallel of the trend for more and more IT infrastructure and services that are provided through smart networks, i.e., cloud computing. The microcomputers are both powerful and autonomous in nature, referred to as embedded systems, and are increasingly being connected by wireless networks, connected both with each other and with the Internet. This is supporting the connection between the virtual and physical world, i.e., the development of CPS (Kagermann et al., 2013).

Even though the storage of data has moved to a large extent from analog to digital storage devices, the content that is stored has not significantly changed during this transformation. Text and motionless images still correspond to the large content of information that is stored in favor of media-rich audio and videos as the general perception about what the digital age is synonymous with. Another interesting aspect of the digital age is whether it is more or better technology that is driving the development of more information. Since the late 1980s, the growth in telecommunication capacity has grown twice as much by the contribution of “better technology” compare to “more technology” and for the explosion of information, the contribution is more than four times from the “better technology” (Hilbert, 2012).

Hilbert and López (2011) have divided the analog and digital technologies into communication, computation, and storage according to Figure 5. The communication of information, or transmission through space, is a measure of the amount of information that is received and sent by a user over a longer distance. One purpose of communication is to overcome transmission over longer distances and hence not the local information sharing. Storage of information, or transmission through time, is the function to maintain the information over a considerable amount of time and to supply the information when needed. To make the meaningful transformation and processing of the information is termed computation. It can also be explained as the action of repeated transmissions of information through space and time with the guidance of an algorithmic procedure (Hilbert and López, 2011).

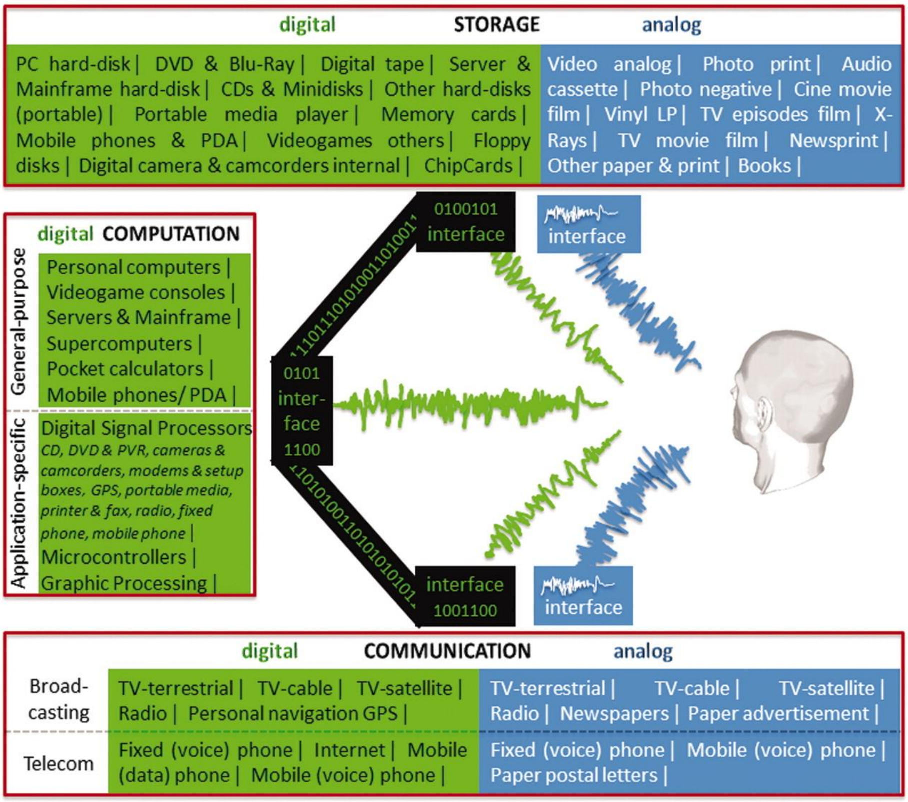

Figure 5. The three basic information operations and their most prominent technologies (Hilbert and López, 2011).

2.5.1 Acquisition of Data

This section will limit the scope of data acquisition to the two cases that are of the relevance of this thesis, namely machine data and spatial data capturing.

Data sources that are already used in the production system has been addressed in section 2.4.3, which are of relevance for the collection of machine data. Besides the secondary sources such as enterprise resource planning (ERP) system, that are important operational data to describe the behavior and condition of the machine, there are new opportunities to collect more IoT and sensory data that are of the primary data source. Sensors react to a change in a physical condition and provide a measurable representation to this change. Sensors are the hardware devices that collect the observed change in an analog signal and can convert it to a digital signal by an analog-to-digital converter. The change of the physical condition includes e.g., temperature, pressure, voltage, current and when the analog signal has been converted to a digital representation, this is transmitted to a processor for further processing (Gungor and Hancke, 2009) and this by wired or wireless networks (Chen et al., 2014). Nowadays, wireless sensors can be used on a large scale to monitor process but also for connecting the entire business end-to-end ("More Than 50 Billion Connected Devices", 2011). As the physical world in the production system becomes more connected, including the humans, machines, devices, and sensors, it is of high importance that common communication standards are used. To have standards in place can support the combination of machines from different machine vendors in a more flexible manner (Gorecky et al., 2014).

3D laser scanning, also called Light Detection and Ranging (LiDAR), is a non-contact method with active sensors that can capture spatial data in three dimensions (Berglund et al., 2016). For capturing spatial information, the methods can be of either tactile or non-contact method. The non-contact method implies that there is no physical interaction with the surface that it is collecting spatial data from and it uses light, sound or magnetic fields instead (Várady et al., 1997). 3D laser scanning technology is using active sensors which imply that it uses a media that can interact with the object of interest for collecting spatial data, which is structured light in this case (Berglund et al., 2016). The structured light is a laser that is emitted along a trajectory from the center of the measuring device and calculates the distance to an object by sampling the reflection of the laser. As the laser beam is swept around repeatedly in the environment that should be digitalized, information about direction and distance measures are saved in the form of a dense 3D point cloud (Berglund et al., 2016; Klein et al., 2012). 3D coordinates are collected automatically in a systematic way and the technology can perform this in the range of one up to hundreds of meters in distance, with hundreds to thousands of points per second (Várady et al., 1997).

2.5.2 Data Storage

The cost to handle data has decreased by big measures at a similar pace as the power to handle it has improved. This development has encouraged business by the economic motivation to keep data stored, since data can be reused for the same or similar purposes (Mayer-Schönberger and Cuiker, 2013). The development of disk technologies has resulted in more commoditized and efficient disks. An example of storage that is using multiple disks is the Hadoop environment, see Table 2 for more information (Davenport, 2014). PCs, handheld devices, mainframes, servers and video game consoles belonging to the general-purpose computers and the application-specific computers, e.g., embedded into electronic devices, household appliances, and monitors, are also examples of technologies of storing devices that has seen a development to support the handling of data (Hilbert, 2012).

Table 2. Big Data technologies in different phases of the data value chain (Davenport, 2014)

Technology	Definition
Hadoop	Open-source software for processing big data across multiple parallel servers
MapReduce	The architectural framework on which Hadoop is based
Scripting languages	Programming languages that work well with big data (e.g., Python, Pig, Hive)
Machine learning	Software for rapidly finding the model that best fits a data set
Visual analytics	Display of analytical results in visual or graphics formats
Natural language processing	Software for analyzing text – frequencies, meanings, etc.
In-memory analytics	Processing big data in computer memory for greater speed

2.5.3 Data Analysis – Transformation to Decision Support

The definition of what the digital age is most often includes the telecommunicating internet and mobile phones, and the large information-storing server farms and databases. The fastest-growing operation of information, however, turns out to be the computation capability. For the general-purpose computers, the computing power has actually grown twice as fast compared to the world's storage and telecommunication capacity. The increase in computing power for the application-specific computers has even grown three times as fast in comparison to storage and telecommunication (Hilbert, 2012). In the big data era, the responsiveness of the communication will be an important aspect for the timeliness of data, and the data can be of either real-time or off-line analysis character. For the real-time data, the analysis and presentation of the analytical results need to be presented in a rapid manner because the data is constantly changing. The offline analysis, on the other hand, is the data analysis that has lower requirements when it comes to response time. Examples of offline analysis are machine learning, statistical analysis, and recommendation algorithms (Chen et al., 2014). Some of the technologies that are being provided with the big data era will make it possible to not only handle the big volumes of unstructured data but also process and analyze it for further usage. They can support for the analysis of textual, video and audio content (Davenport, 2014) and an overview of the key technologies that are mentioned for handling big data are summarized in Table 2.

2.5.4 Transmission of Data

Transmission of data is the functionality to exchange data over a longer distance and the connectivity solution is an important aspect for the realization. A network has the function of transmitting data from one node to another (Moyne and Tilbury, 2007). From a technical definition, a networked control system includes the system to be controlled, actuators, sensors, and controllers. These are controlled by a communication network. Previously, a control system connected the components with each other through hardwired connections and the data collected at sensors were sent to a central location where it was used for monitoring. It constituted a basis for making decisions about how to act. Nowadays, however, with the technology available, low-cost processing power can be installed at a remote location via microprocessors and the data can be transmitted in the shared digital networks or even through wireless connections with high reliability (Baillieul and Antsaklis, 2007). The traditional point-to-point connections with a wired solution can now be replaced by a control network that brings a number of advantages.

Industrial automation systems have been realized through wired communications and this requires installation and maintenance of expensive communications cables. Because of the high costs involved this has been a cause for the limited implementation in industries ("Assessment study on sensors and automation in the industries of the future", 2004). The wireless sensor networks (WSNs) has seen advancements and made the realization of embedded industrial automation systems feasible to a lower cost (Akyildis et al., 2007). Besides the lower cost involved with this technical solution, other advantages that the industrial wireless networks (IWSNs) can provide are self-organizing, rapid deployment, flexibility, and inherent intelligent-processing capability. In the systems where the wireless solution is used, sensor nodes can be installed on the industrial equipment and thereby control the parameters that are critical for the equipment (Gungor and Hancke, 2009). One difference between two wireless solutions, the WLAN/WiFi and the wide areas wireless access solution, such as 3G and LTE, is the distances related to the access point that the connection can be provided. For a WLAN it is normally within a few tens of meters from the access point and for 3G and LTE a few tens of kilometers from the base station (Kurose and Ross, 2013). One of the most obvious advantages with wireless networks is the reduced volume for wiring that otherwise is needed. Another is the reduced numbers of physical points, such as connectors and wire harness, in the control network that can cause failure for the entire system. This brings the positive effect of creating a more reliable system. More devices and units, ranging from a broad variety of applications, are becoming connected and demands of constant availability, resilience, coverage, latency, and bandwidth will be of importance ("More Than 50 Billion Connected Devices", 2011). The next generation of mobile communication, 5G, has the potential to serve these needs (Osseiran et al., 2014). Technical requirements for 5G will be a system that is capable of 1000 times higher mobile data volume per area, 10 to 100 times higher number of connected devices, 10 to 100 times higher user data rate, 10 times longer life of battery (for low power massive machine communications), and 5 times reduced End-to-End latency (Osserian et al., 2013).

2.6 DECISION SUPPORT

Making a decision is the action of doing a selection in a situation that could not otherwise be resolved. In a decision-making process, data is providing the support for making an informed decision and it can be seen as the tool for the decision maker (Košcielniak and Puto, 2015). Information that is used for the purpose of handling an event, can be described as data that is used and presented in a specific manner at the appropriate time that can improve the knowledge of the receiver for taking a specific action or making a specific decision (Galliers, 1987). How information should be presented can be defined by the content and carrier. The carrier implies how information is presented, i.e., the medium presenting the information. The content is what information that is presented, i.e., the mode of presenting the information (Fässberg et al., 2012). Information that will support in event handling, can be categorized into three different types; the problem, domain, and problem-solving information (Byström, 1999; Byström and Järvelin, 1995). When it comes to the information related to the problem, this should contain information that can represent the specific event that has occurred and needs to be handled. This involves the description of the structure, properties, and requirements of the problem. The domain information includes known facts, concepts, as well as theories relevant to the domain of the problem. When it comes to the information that will support the persons that will solve a problem, it comes to the category of problem-solving information. This information will rely on both the problem information and the domain information in order to guide the person to cope with the problem at hand (Byström and Järvelin, 1995).

Information is further processing of the data that has been collected to bring support in the decision-making process, but what distinguish information from knowledge? Because even though they are often used in the same context, there is a distinction between them. Information can add, restructure or change the knowledge one has about something and this is done by the flow of message or meanings that the information brings (Machlup, 1983). So if the information is the actual flow of messages, knowledge is the instant that is created and organized by the messages in a combination of the commitments and beliefs that the receiver holds. For creating knowledge, the information as a medium or material is, therefore, a necessary piece.

Information can be viewed from two perspectives; the syntactic and semantic perspective. The syntactic aspect is what has been discussed earlier in this chapter, the volume of information. Since it concerns the quantitative measures of information, it excludes the actual meaning or value of the information. The semantic view, on the other hand, is the consideration of the new meaning of the information piece (Nonaka, 1994). An example of how to describe the semantic meaning of information has been expressed by Bateson (1979) as *"information consists of differences that make a difference"*.

Hermann et al. (2016) foresee a change in responsibility of humans in the production systems as a result of the realization of the Smart Factories of Industrie 4.0. The role of humans will shift from operating the machines toward a role that incorporates more strategic decision-making as well as being a flexible problem solver. The complexity of the production systems increases with for instance the full realization of the CPS and this will require assistance systems that can aggregate and visualize information to the humans to empower informed decisions as well as to handle the urgent problems that require actions on short notice (Gorecky et al., 2014). Smartphones and tablets are an example of the carrier that will play a central for connecting humans in the production system with the IoT (Hermann et al., 2016).

3

RESEARCH APPROACH

Research is the process of pursuing a systematic investigation, which involves to concentrate the thinking as conscious efforts and to do it in a rational, careful manner. A fundamental aspect of the research process is that it should be conducted in a way that can enable people to understand, reproduce, and evaluate the quality of it (Trochim et al., 2016). The systematic collection and interpretation of data and the clear purpose of why data is collected are key characteristics of the research process. The systematic way that research is carried out leads to findings of the problem addressed and thereby also contributing to new knowledge (Saunders et al., 2016). This chapter explains the systematic way this research has been performed, what worldview and research approach it relies on, and how the research questions have been answered with the research projects, studies, and appended papers. The chapter will go through the philosophical worldview, continuing with the approach, design, and data collection methods used.

3.1 THEORY BUILDING

For the reasoning and contribution to theory, there are mainly two different branches called the deductive and inductive reasoning. Inductive reasoning implies to start from the specific observation and move to the general theory building. It is more open-ended and exploratory in its character (Trochim et al., 2016). Theory built from the inductive reasoning always starts from the observations made hence the data collected, and a researcher using induction is therefore interested in exploring the phenomena in the context where it takes place (Saunders et al., 2016). Deductive on the other hand starts from the general and moves to the more specific. Therefore, it is narrower in its nature and the main activity is to test or confirm posed hypotheses (Trochim et al., 2016). Deductive reasoning is when conclusions have been derived in a logical manner from the premises that have been set for the study and that the conclusion is defined true when all the premises are true as well (Ketokivi and Mantere, 2010).

This research was first initiated from a phenomenon that was studied, namely that the current practice of data was scarce in the manufacturing domain. This triggered a need for more investigation in this area, also supported by the already available literature in the research community stating the importance of this problem, especially the absence of data representing the reality in the simulation case. As the research in this area was continued, the concept of data analytics and now also big data were identified as relevant areas for investigation in the real-world context. This has motivated the other studies in this thesis, to go from the more general to investigate it in the specific case and make the findings generalizable. In one of the studies, it started from the more general, examined it in a specific case with an interview study, but the results allowed to provide new insights in the area moving to the more general again. The research in this thesis can be explained as moving back and forth between the deductive and inductive reasoning.

3.2 PHILOSOPHICAL WORLDVIEW

When pursuing research, the justification of how it is done with the choices of methods and methodologies is something that is based on the assumptions that the researcher makes about reality (Crotty, 1998). To explain the philosophical worldview implies for the researcher to explain the general philosophical orientation and the nature of the research that the researcher brings to the study (Creswell, 2014). To acknowledge the research philosophy is to explain the system of beliefs and assumptions made about how knowledge is developed. These assumptions are fundamental for how the researcher understands the research questions, the methods that are used and how the findings are interpreted (Crotty, 1998). For the formulation of a research problem, one common source to use is the existing literature in a field. The experience of a practical problem identified in a field can, however, be equally important. For applied research, a great source for motivation for the researcher can be the opportunity to make a difference to a problem that has personal importance (Trochim et al., 2016). The intended stakeholder for this research is the manufacturing domain and the research itself is applied oriented. The background and previous experience in the manufacturing domain by the researcher are therefore acknowledged. Prior to initiating the Ph.D. studies, experiences have been gained from both studying and working in the field of manufacturing in Sweden and in an international context, Germany and USA. The undergraduate studies were carried out in Mechanical Engineering and the work experience was from an engineering company acting at an international arena with production sites all over the world. The national and international experiences from working in industry have provided a lot of insights into the challenges in industry. Both from a technical perspective, but foremost from an organizational. These experiences arose the curiosity to provide more knowledge of what the future factory can look like and that is attempted to be addressed in the research.

The assumptions about what creates acceptable, valid and legitimate knowledge are referred to as the epistemology (Burell and Morgan, 1979). One of the epistemologies, presented by Creswell (2014) is the pragmatism. Pragmatism is focused on the consequences of actions and concepts are only relevant where they can support the actions (Kelemen and Rumens, 2008). For the pragmatic worldview, the research is initiated by an identified problem and the research that is being pursued intends to identify solutions that can support in future practice in the field (Saunders et al., 2016). The research is centered and outlined around what best can support the research problem and hence the research questions raised. For the pragmatic worldview, the consideration of what works to solve a problem is the main objective and emphasizes, therefore, the importance of focusing on the research problem and select the methods that are the most suitable in order to address it (Creswell, 2014). As was mentioned, the research performed in this field is to a large extent application oriented, where it is investigated and explored how new technologies and concepts can provide value to the manufacturing domain. Data is used directly from the real world application and with the knowledge gained it can be generalized for the research community. The pragmatic theory is, therefore, an appropriate way of explaining the research field saying that a statement is true if it can provide value when put into practice. The main focus is to address the problem identified and whatever method that is regarded as suitable is motivated to be used. For this licentiate thesis, the problem area is explored by describing, explaining and suggesting, hence having an explanatory goal of the research. The mixed methods research that combines the qualitative and quantitative research approaches further explained in section 3.3, has the benefit of looking at numerous ways of how to collect data in order to answer the problem rather than subscribing only to one of them (Creswell, 2014).

3.3 RESEARCH APPROACH – MIXED METHODS RESEARCH

For the research process, a general plan is required for identifying how the research questions should be addressed and this is achieved when the research approach is decided on. The research approach will either be of quantitative, qualitative, or mixed methods research character. The qualitative research is the approach that collects the more open-ended data and the intention to explore and understand the meaning of individuals or groups that can be ascribed to a social or human problem. The quantitative research is the approach for testing the objective theories, which is done by the examination of relationships between identified variables and the data collected are close-ended. Since the quantitative and qualitative research approach both encompasses benefits and weaknesses for answering research questions, mixed methods research combine both in order to gain the strengths and minimizing the limitations of them. The core of this approach is that a more complete understanding of the research problem can be gained and it is aligned to the pragmatic view (Creswell, 2014).

The research in this thesis has applied the mixed methods research in order to gain a more in-depth understanding of the problem, see Figure 6. Depending on scope and question raised in the studies, the data collection has varied between qualitative and quantitative studies. Most studies included in this research has been initiated by quantitative data collection, collecting production data for a simulation project, connecting a grinding machine to enable more machine data or applying 3D scanning for capturing the spatial data. To explore the benefits of having more data available but also to understand the current practice of data, has motivated a more qualitative approach. This is because it has not been possible to measure quantitatively in the studies performed and also because an important aspect of the thesis is to explore the benefits the organizations can recognize by the use of more data. In order to answer this, the qualitative data collection has therefore been used for answering the questions. Creswell (2014) called this approach the explanatory sequential mixed methods, where quantitative data is first collected to be supplemented by qualitative data.

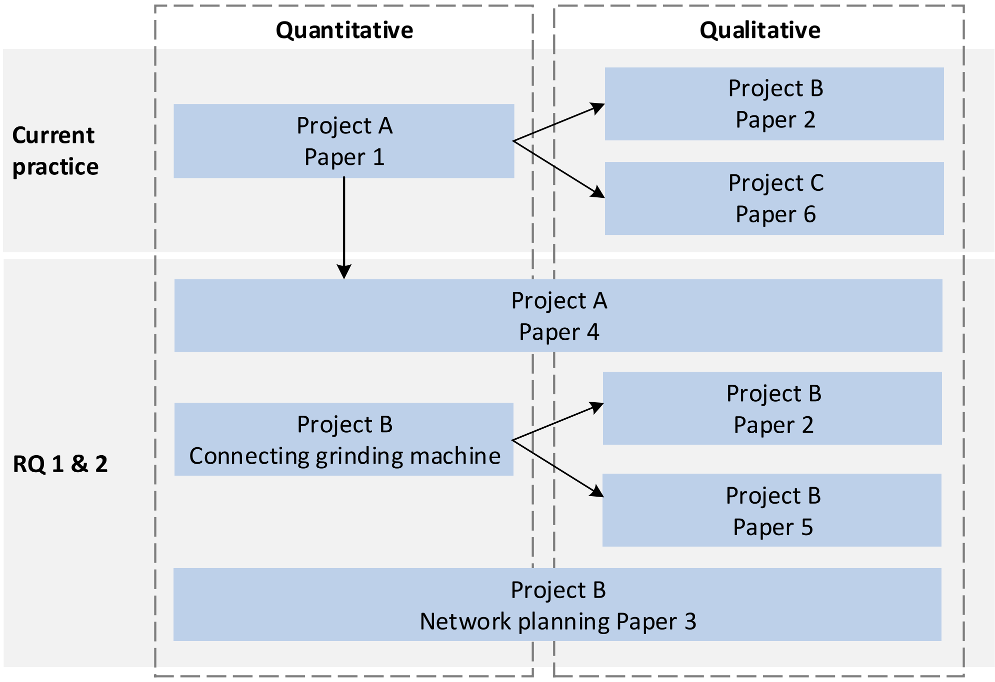

Figure 6. The mixed methods research approach applied in this thesis

Three research projects have served to answer the research questions. They will be explained in this section to provide the required knowledge about them before continuing to explain how data was collected from each of them based on the mixed methods research approach. The timeline when they were performed and the belonging paper is visualized in Figure 7.

Project A is a joint research collaboration between RUAG Space Group AB, Sweden (RUAG) and Chalmers called "The space industry of tomorrow". RUAG is a contender on the communication satellites market which is experiencing market changes and the project has therefore been defined in three deliverables; a current state analysis, a high volume scenario, and a flexible volume scenario for the production system. Paper 1 presents the first step of the study, collecting data from the production system for the purpose of a simulation project. Value Stream Mapping (VSM) was identified as a suitable data collection strategy and involves both analyzing secondary data, observations in the production system, and interviews. It was suggested from the literature on how the data availability should be categorized, which was also part of the results of this paper. The second part of this project is presented in paper 4, using the results from paper 1 as an input, where a discrete event simulation (DES) model is built. 3D laser scanning was used as a supporting tool to capture the spatial data. The virtual environment was presented to project members and other personnel knowledgeable of the production system to validate the simulation model. Focus group interviews were used for the evaluation of how accurate the virtual representation corresponds to the real-world system.

Project B is a testbed project named 5G Enabled Manufacturing (5GEM), a collaboration between Chalmers, SKF, and Ericsson AB. The main idea is to investigate how the next generation of telecommunication, i.e., 5G can be applied to enable higher efficiency, flexibility, traceability, and sustainability in the manufacturing domain. All three papers written within the scope of project B, rely on the same initial efforts. This involves deploying a cellular radio network in the workshop, connect a grinding machine, enable the data exchange and develop appropriate algorithms for the identification of correlations in data, and send the data back to a mobile operator support system. This is the design of the experiment where the system is changed in order to evaluate new technology. The second part of paper 2 and 5 further explore how the availability of more data can provide benefits in their organization, both the data itself and as decision support, which has required qualitative methods. Paper 3 is focused on testing a developed offline tool that would support the future installation of the cellular network done in this study, which was performed by first quantitative collection and complemented with observations made by the participants for evaluating the offline tool.

Project C is a multi-case interview project with two case companies from two different branches. One of the case companies is RUAG Space Group, from project A and the other case company is from the robotics industry. The aim of the project was to evaluate standardization challenges experienced with data acquisition and it is in this thesis supporting to describe the current practice of data. To receive more diverse answers to this topic, two case companies with different procedures and routines for data acquisition for the simulation purpose were selected, i.e., a multi-case study and a qualitative approach was used.

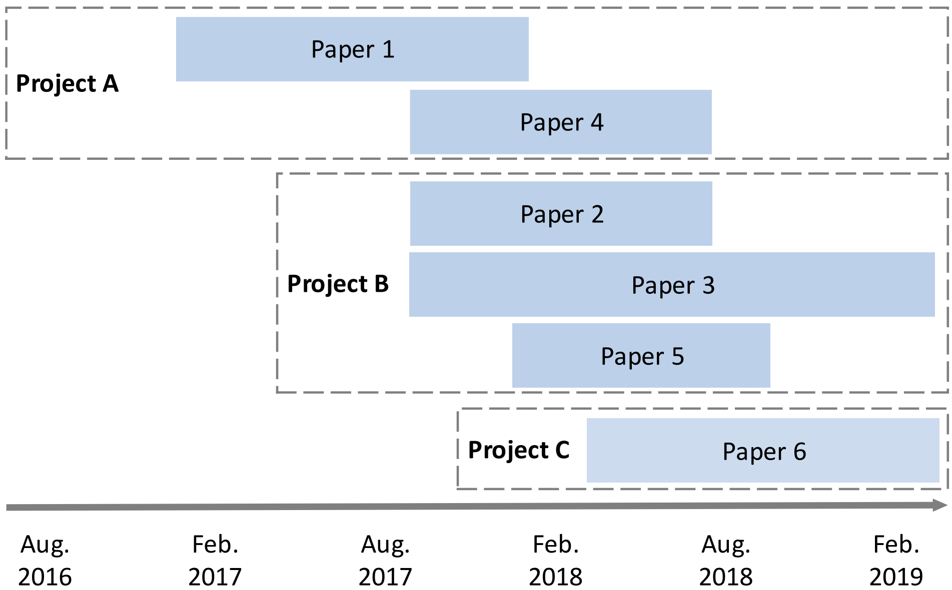

Figure 7. Summary of studies, with belonging papers and timeliness

3.3.1 Data Collection Methods Used for Answering the RQs

How the mixed methods research was used to answer the research questions is illustrated in Figure 6 and the specific design and data collection methods are summarized in Table 3. It should be noted that the literature review has been used throughout the research and is therefore not specifically mentioned in Figure 6 or in Table 3.

The current practice of data
Data from the production system for a simulation project were collected at an initial stage. This endeavor revealed that the data available for this production system was scarce and hence suggested that the practice of data for this production system was low. The results provided insight into the three categories A, B, and C of data availability presented in section 2.4.2. Based on the results, the topic needed a more in-depth examination to raise questions about the current practice of data. Project B, paper 2, was performed in the context of a focus interview with production personnel evaluating the result of the technical implementation done. The questions raised were formulated based on literature. To understand the expected benefits of this area, the current practice was also evaluated in the focus interview. Project C, paper 6, was performed in order to explore the current practice of data based on the trust and availability of data in a semi-structured interview study. It involved a multi-case project research design, which allowed a benchmarking between two different companies with different practice of handling data. In each case company, a number of interviewees were identified that have a different usage of the production data in their daily operations. The interview questions were developed by the challenges in data acquisition for simulation projects stressed in literature. This project contributes to an empirical identification from a real-world context to what has been expressed in literature.

RQ 1: From reality to data and *RQ 2: From data to decision-support*
For paper 2, project B, and paper 4, project A, it was identified that interviews with subject matters of the organizations could provide the insights needed to receive a more comprehensive understanding of the research problems. This was performed in a sequential way, first collecting field data with the help of digital technologies and in the second step evaluating the benefits with the help of interviews. Paper 2 involved a focus interview with participants that can be regarded as receivers of the data in the mobile operator support system, and the questions raised were inspired from the literature explaining the big data concept. Paper 4 also includes a focus group interview with personnel knowledgeable of the system for evaluating a new design.

Paper 5 takes a different perspective compared to paper 2 and 4 since the main aim is to explain the knowledge and learnings that have been identified from the first quantitative step of the study. Instead of performing an interview to receive a more in-depth understanding, observations have been made by the researchers to evaluate the potential benefits of having more data available enabled by digital technologies. The researchers have mainly been gathering field notes from observations that have been conducted as participants of the study. Documentation in the form of meeting protocols and documentation of the different steps of the project have been recorded and also used for the analysis in order to answer the research questions. This provided answers to the potential benefits.

Paper 3, project B, is a joint effort to investigate a new way of working where both the expertise of the manufacturing and radio propagation fields have been combined. An initial step was to collect spatial information of the factory environment that is difficult to gain otherwise without the digital technology. This first step is providing the environment for the second step where the simulation is performed. The simulation results are compared to the real measurements of the network and this could indicate the applicability of the new offline tool that has been developed. The qualitative evaluation was performed along the project by observations made by the participants, but this time the participants where gathering field notes spending more time as a participant than as an observer (Creswell, 2014).

Table 3. Summary of the papers with selected research design and methods for data collection

Paper	Research design	Data collection methods
1	Field study	Observations, interviews, and secondary data collection
2	Interviews	Focus interview
3	Experiment Observations	Measurements and simulation Field notes
4	Interviews	Focus interviews
5	Field study and observations	Field notes by observations as a participant
6	Multi-case study	Semi-structured interviews

3.4 RELIABILITY AND QUALITY IN RESEARCH

An important aspect to be considered for the research design is how this is done with respect to ethical issues and possible constraints that will be encountered in the research (Saunders et al., 2016). Besides the ethical issues and constraints, the research also aims to gain results that are useful for answering the questions raised, reach meaningful conclusions and in overall a deeper understanding and advancement in the area rather than seeking the absolute "truth" about a research question (Trochim et al., 2016). What identifies research to be of high quality between qualitative and quantitative research. Credibility for qualitative research requires the confirmation from the participants since they are the ones that can judge if the results are believable. Whether results can be seen as generalizable or transferable to another context for a qualitative study tells the transferability of the research. From the quantitative area, the reliability of a study has been a measure to tell how repeatable the research is. For the qualitative studies, this is referred to as the dependability of the study and means to what degree a phenomenon can be observed twice. Since the context can change, the researcher is also responsible to describe the changes and how they may impact on the results presented. Confirmability of the results and the conclusions made is a quality measure to reduce the potential bias a researcher brings to the study. There are a number of strategies available to increase the confirmability, such as triangulation and performing a data audit for examining the data collection and analysis procedure (Trochim et al., 2016). The reliability and quality of the research presented will be discussed in chapter 6 Discussion.

4

RESULTS

This chapter presents the findings from the appended paper that support to answer the research questions. This is why the chapter is structured according to the research questions. Firstly, results are presented that show the current practice of data. Secondly, research question 1 and 2 are addressed respectively with a table summarizing the values identified.

4.1 CURRENT PRACTICE OF DATA IN PRODUCTION SYSTEMS

To understand the need for more data to drive decision-making, the current practice of data in the production system need to be described. This section will exemplify and explain what the availability of data but also the challenges involved for data acquisition in the production system are for stating the current situation. This will support the following sections where the delta between before and after applying digital technologies to make more data available will be presented.

For the purpose of building a simulation model of a production system, paper 1, an approach was designed for the collection of data. It combined efforts from Value Stream Mapping (VSM), Improved VSM (IVSM) and Value Network Mapping (VNM). The production organization has an engineer-to-order strategy and the production system is characterized by complex and merging product flows. The awareness of what data could be collected or not for representing the system was limited. A last step of the approach was therefore included, categorizing data parameters to A, B, and C (see section 2.4.2). This was an important step of the approach in order to understand the availability of data for this production system and to identify further actions for collecting these. While performing the mapping activities of VSM, questions were raised whether data parameters of interest is already available, collectable or can be estimated. The result of these actions are summarized in Table 4 and as can be seen, all three categories are represented in this system.

Table 4. Classified data in category A, B, or C and source for data from paper 1

Data parameter	A	B	C	Data source
Process time	X			ERP-system
Cycle time	X			ERP-system; process time/batch size
Change over time		X		Measure with stopwatch or estimate
Batch size	X			ERP-system/Manufacturing order
No. of operators	X			Data from the manager for the production segment
Scrap rate			X	Non-conformance report (NCR)
WIP			X	Manually count WIP in the value stream
Transport distance		X		Measure in either 2D-drawing or physically in the workshop

Category A was data that are documented in the ERP-system such as batch size, process and cycle time but also the number of operators that the managers for each production unit have documented. Change over time for resources and the transportation distances can be collected, but they have not been documented so they belong to the category B. Data parameters that have a big impact on the quality yield and performance of the system were neither available or collectable hence of category C. The data source for scrap rate is the NCR which means for the production organization that there is no performance measure for the quality yield of their production. To identify the quality yield they will have to extract the cost involved for each product individually that have been documented in the NCR, because that it is the only way it is documented. They neither have a performance measure that can inform about the tied up capital in WIP, required storage or space for each production area since the products are hiding in storages in the production environment that need to be estimated by manual collection to get an indication for this number.

The availability of data in the previous case could be determined by the effort made when collecting data for a simulation project. In a later study, paper 2, the current practice of data was explored by addressing questions about how decisions are performed today, based on data or intuition, and the three aspects of big data, i.e., volume, velocity, and variety of data. The topics raised and the answers received are summarized in Table 5. On the question about how decision-making is done today, it was recognized that it is mainly the intuition and experience by personnel that drives the decision-making. There is a become more to data-driven decision-making and it is a reason why the maintenance organization has started to collect more data to gain an understanding of the behavior and condition of the equipment in the system. For the three aspects of big data, it was discussed how each of the aspects are important for supporting decisions. The volume of data was identified as important for the production technicians and the maintenance personnel since they are looking at the behavior of the system over a longer time period to identify trends. The velocity of data was considered more important for the daily operations since it requires to receive data fast to make the required actions. For the variety of data, it was recognized that it is scarce in the current system. For sharing information between the shift, there is one system where data is added manually and there is also an oral communication for exchanging information between personnel.

Table 5. Current use of data and the importance of volume, velocity, and variety of data from paper 2

Question	Current state
Data or intuition	Mainly based on intuition. Maintenance has started to collect more data.
Volume	For production technicians and maintenance, it is the volume that will play an important part to drive decision-making.
Velocity	It is important for daily decision-making.
Variety	Data sources used today is a system where the operators manually add information about status when the shift is finishing. There is also oral communication between the personnel.

The challenges stated are accuracy, correctness, duplication, consistency, timeliness, validity, reliability, and completeness of data. The results from the interviews at two different case companies, paper 6, are presented in Table 6 and the interviewees were asked about how the challenges impact on their ability to use data for driving decision-making.

Table 6. Challenges of data acquisition found in Case A and B from paper 6

Challenges of data acquisition	Case A	Case B
Accuracy	Uncertainty about what times that should be used and the data is not accurate with many factors impacting.	Free from errors, machine data are automatically collected and the manual collection is performed in a systematic way.
Correctness	Different labeling for the same operation because of no standard way of reporting. Pre-set values can be used in ERP-system instead of feeding actual times. Data need to be adjusted and compared to statistics.	Not analyzed the data yet but it is not considered as an obstacle in this case.
Duplication	Multiple fields in the system are used and the same data item is stored in both ERP-system and spreadsheets.	Has not been encountered in this case.
Consistency	Differences between departments of how times are reported in ERP-system and different results depending on the data source.	Measurement repeatability and reliability is an issue for manual collection.
Timeliness	Change to the system impact on data over time. Data changes in reality but not in the system.	Is a challenge when collecting and sampling data.
Validity	Not describing the behavior of the system, discrepancy between reality and data in the system.	Data collected matches the actual processes and has been validated.
Reliability	Unrealistic data has been encountered and there is no common understanding within the organization of how to use data. Planning is always questioned by other departments.	Has not yet been evaluated, but the data collectors are trusted.
Completeness	Do not provide all necessary parts and assumptions need to be made with gut feeling leading the way. Estimations of the data are done to get more accurate planning.	Has not been encountered as an issue for this project yet, but is expected to be later on.

What can be seen from the results of the two case companies, is that the challenges are acknowledged by one of the companies to a larger degree. The companies have different routines of handling data, a reason why they were picked for this study, where one of them is collecting data mainly by manual efforts and the other company has an automated data collection for processes in the production system. The company where the data collection is performed manually could recognize issues for all eight aspects presented in Table 6. When using data for decision-making, they have uncertainty in the data, the correctness of the data is questioned because it is labeled differently and pre-set values can be used instead of actual outcome, and data becomes duplicated because it is stored in multiple sources. Because of the lack of quality for these aspects, it can be seen that it also impacts how valid, reliable, and complete the data is considered. This is questioned by the interviewees themselves and by the rest of the organization as well. The other company where the data collection is more automated do not recognize these challenges to the same degree, or it has not yet been encountered for some of the aspects, but there were three areas where they agreed. Consistency, completeness, and timeliness of data are issues, especially when data has been recorded by manual efforts. For the case company A, it was mentioned that the data may change in the reality as a result of changes made to the system, but it is not necessarily changed in the computer system where data is stored since pre-set values can be used for reporting an operation.

4.1.1 Why Decision-Making Should Be Driven by Data

In the interview study, where challenges with data acquisition were explored, the interviews were of a semi-structured character which allowed to go further into some of the areas that were identified as interesting, paper 6. The results are presented in Table 7 and the topics addressed where the connection between data, information and decisions, ownership of data, and alignment to performance measures.

Table 7. Representative quotes of how data is identified as valuable for driving decision-making from paper 6

Case	Representative quote
Establishing a connection between data, information, and decisions from a holistic perspective	
A	"*Not knowing what data that contributes to the lead time is a consequence of a lack of a holistic perspective in the organization. Data helps to solve this issue. I can understand our processes better through data.*"
B	"*The lack of a holistic perspective is a consequence of the dispute when people have different backgrounds and functions. It is natural to perceive things differently. This is an issue that reflects in our simulation.*"
Ownership of data across different functions in the production process	
A	"*Data comes from the activities that go on at the factory floor. It is important for the operators to report data. The system cannot help us as long as we are not responsible for feeding the right data to it.*"
B	"*There is a challenge in the tendency to invest a lot of technology and time into collecting all sorts of data, having huge databases, and not doing anything with it, or answering very basic questions with it.''*
Alignment of data to performance measurements leading to improved decisions	
A	"*Data acquisition is a top-down effort which needs the interest of management to follow up. Performance measures are needed for this. We need to measure to decide on something, and measurements need data.*"
B	"*That data is needed, is not visible if there are no measurements. More importantly, we need to understand why we are measuring. Because if there is no value in our measurements we cannot justify the cost of acquiring data*"

4.2 RQ1: FROM REALITY TO DATA

This section will continue to present the results received in the studies where digital technologies have been applied for collecting more data. The section will describe the delta of before and after the implementation of digital technologies for three different areas; a connected machine, building a simulation model, and factory radio design. The results will support to answer how data were digitalized by digital technologies and sequentially how data were found to bring values to the production organization.

4.2.1 Connected Machine

In the context of project B, it was explored what the possibilities are in a factory with unlimited connectivity are. The initial stage was an NC machine with a machine computer. Data is collected by existing sensors to provide the human operator with the status of the machine and if any actions are required. The data can be accessed by the Human Machine Interface (HMI) at the machine, but the data is not extracted from the machine or used for performing any trend analysis or evaluating the historically collected data. After machining, each machined part is measured by a stand-alone measuring machine to identify any tolerance offset. There are mainly three machine components that are critical for the performance of the machine, namely the ball screw, slides, and motor. As these are worn out, they may impact on the quality of the product implying re-work or even scrap of the product. To collect data parameters for these three components is, therefore, consider of high importance.

To explore how a connected machine could provide values to the production organization in decision-making, Ericsson AB provided the connectivity infrastructure and ICT expertise. SKF provided the manufacturing environment and manufacturing domain expertise. The technical setup is fundamental for three of the appended papers and therefore explained in this section to provide a basis for the results. A dedicated 4G LTE network was deployed in the workshop and an overview of how the machine was connected is shown in Figure 8. The machine and mobile units were connected to a cellular LTE network exploiting LTE modems. The distributed IoT platform Calvin is the platform used for data collection

(Persson and Angelsmark, 2015) and there is a local client server in the workshop for intermediate storage. Data are sent to where the database is geographically located, ca. 250 km away in Lund, and stored in the cloud. The analytics platform, which holds the analytics capabilities, is geographically located ca. 450 km away in Stockholm and performs the analysis and anomalies detection.

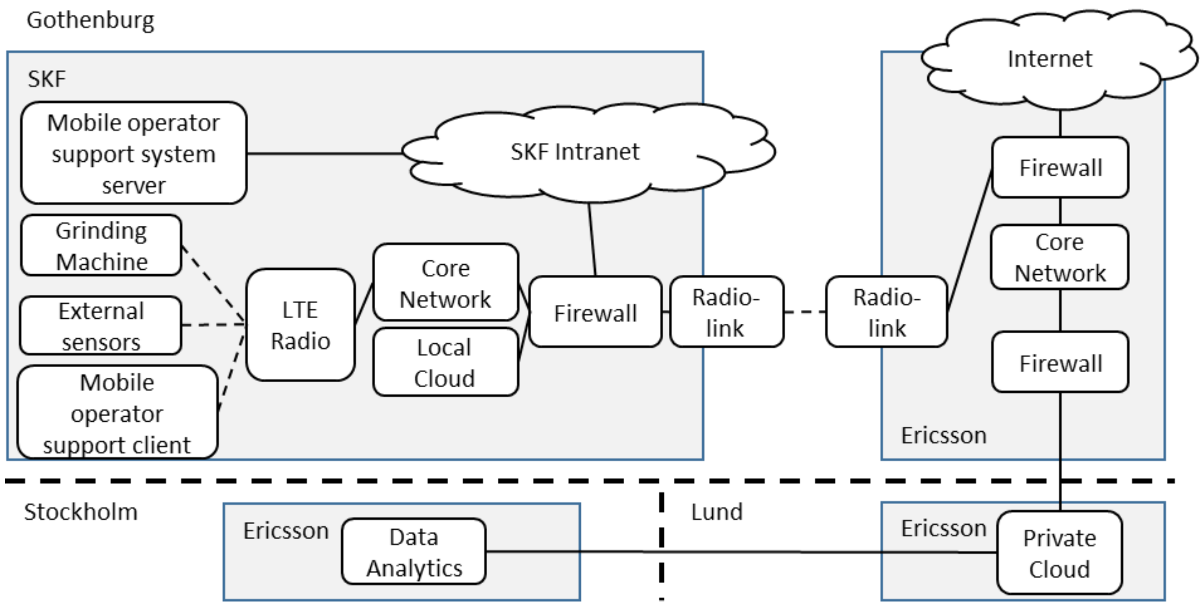

Figure 8. The infrastructure of the connected machine for the 5G testbed project

External sensors and Programmable Logical Controller (PLC) data from the machine were part of the scope for collecting data, see Figure 8. As was mentioned, there are three components in the grinding machine that are critical for the performance of the machine (the ball screw, slides, and motor) and the technical setup was enabled to collect data about these. To satisfy two different approaches of data collection, four different ways of collecting data were used and could be divided into the two main groups internal machine parameters and externally mounted sensors, which are further explained in Table 8.

Table 8. Data parameter of interest for the connected machine divided into internal and external data parameters

Data parameter	Description
Internal machine parameters	Sensor signals and program systems parameters accessed from the internal machine computer. In total 43 individual tags were possible to collect from the machine computer
Externally mounted sensors	A vibration measuring system called IMx, two externally mounted sensors measuring cooling fluid and temperature connected with the communication protocol IO-link, and a temperature sensor embedded on an RPI.

Data were collected every 100 ms and for the data transfer, the machine computer in the machine and IMx both supported the standard communication protocol OPC UA and was therefore used. The IO-link sensors send data as a Transmission Control Protocol (TCP) data stream through the IO-link master and the temperature sensor using Raspberry Pi (RPI) can locally be read in the RPI computer and sent by any communication system supporting the Linux platform. To send data to the database, the IoT platform Calvin was used.

There were a number of functionalities that have been developed for this project in the mobile operator support system (MOST) that present real-time data as decision support and three of them are presented here; tolerance offset, vibration data, and cycle time. Real-time machine data are presented to the machine operator in the mobile operator support system client, which will be further explained in the next section answering RQ2. The pre-processing and storage of data was handled in the database in Lund. Different types of data can be handled in the data-center because the document database (Rajkumar et al., 2010) is a noSQL (Renwick and Babson, 1985) database. The Apache Hadoop (OpenCV, 2018) and Kafka (Esmaeilian et al., 2016) platforms are used for distributing the data. Data are stored as JavaScript Object Notation (JSON). For data analytics, anomaly detection in vibration data was the purpose of the data analytics step and Twitter's Luminol library was used (Luminol, 2018). It is suited for continuous and seasonal data. Eight different sensors provided vibration data and anomalies were identified for all of them varying between 5-50 during a working day. For the Root Cause Analysis (RCA), data was also used from torque and temperature, and a Decision Tree Algorithm was used to find the combination of values from the different data. The last step of the data value chain, the system feedback, is presented in the next section answering RQ2.

To understand the applicability and benefits the organization can recognize from having more data available, this was addressed in the same way as described in section 4.1.1 from paper 2. The evaluation was based on the definition found in the literature for big data; volume, velocity, and variety of data (the three Vs of data). Table 9 presents the results from the evaluation of volume, velocity, and variety of data for the future state.

Table 9. Use of data in the future state of the production system with consideration to volume, velocity, and variety of data from paper 2

Question	Future state
Data or intuition	More data-driven decision-making.
Volume	More data will become increasingly important as the organization learns more from the data.
Velocity	For the operator, the speed of data will be important.
Variety	Variety of data sources will be important in the future, to extract and analyze data from various sources for analysis and decision-making.

A comment during the interview was the importance of receiving the right data with the right resolution to support decision-making. Data presented with a resolution that does not corresponds to the needed one on, e.g., cycle times may lead to wrong conclusions or not support the intended decision. Another aspect that was discussed with the focus group was if there is any trade-off between speed and quality of data. The answer was that it depends on the situation and the purpose of the data. For some situations the awareness of that something has happened can be more important than the richness of the data. Hence the speed is more important. But if the purpose is to do an analysis or to identify the root cause of a problem, then the quality of data is more important than receiving the data fast.

4.2.2 Virtual Representation of the Production System

The limitations and opportunities in the facility where the production is taking place are important to consider. Both from the perspective of how space can be used in the best way but also from the human perspective that is working in the environment. To plan a layout of an area already in use or a new space in a production system has traditionally and still is to a large extent based on the information available in a 2D drawing. To capture the spatial in a digital form in order to make it available in a virtual representation has the potential to support in this matter. This section will present the results of two different cases with two different purposes of the application of spatial data. The first is for building a simulation model of an already existing production system and the second is for planning a radio network for a factory environment.

The potential of supporting DES with VSM and 3D laser scanning during the different phases of a simulation project was investigated in one of the studies, paper 4. For this study, the FARO Focus 3D was used to gather data via a total of over 50 scans and a resulting point cloud of over 500 million points. There was no other activity in the production system during the scanning and it took about seven hours to complete the scanning and another six hours for processing the data. The point cloud was imported to the simulation software for providing the virtual environment. The simulation model was built in the software Visual Components 4.0 Premium. In the first step the data gathered through the VSM, already explained in section 4.1.1, was supporting to build the individual processes and connect them to illustrate the entire production system. In the second step, where the initial model was combined with the point cloud emitted by the 3D laser scanning technology, could provide the accurate visualization of the factory. The production processes could then be positioned in the corresponding location of the real word system mimicking the movement of operators and product flows. As this was accomplished an informant from the company could support in guiding where the different steps physically occurred.

The point cloud enabled by 3D laser scanning was easy to import to the simulation software and made it possible to map where the production processes occur in the real-world system without interrupting the production and it also has the benefit of being stored digitally. The informant from the case company could in the 3D representation of the factory easily recognize and point out where each process is taking place since the point cloud was accurate within less than a centimeter and photorealistic of the real-world system, see Figure 9. 3D laser scanning captures the production facility as-is and allows offline work. It also supplements VSM with information regarding the number of machines and workstations.

Figure 9. Point cloud of the workshop imported to the simulation software from paper 4

Connectivity providers need to identify more efficient work methods of deploying the network in the factory environment ensuring a reliable and available network. Within the manufacturing domain, 3D laser scanning has become a commonly applied tool for capturing the spatial information of a facility. There is a big potential of applying 3D laser scanning in order to gain the spatial information of a factory with information about the position of machines and information about the material that impact the scattering effects to do a radio space design offline. An approach including 3D laser scanning and ray-tracing simulation to develop a network plan for installation was developed and evaluated in paper 3. The approach includes four steps where the three first involve building the 3D environment of the workshop and the last step the ray-tracing, physical optics, to receive a network plan. For answering RQ1, the procedure of creating the virtual representation and the results of these steps will be presented. In the approach, this is referred to as Step 1-3: Spatial Data Collection.

Step 1 involves performing the scanning of the environment that is intended to be digitalized. In this setup, the 3D imaging data collection was conducted using a FARO Focus 3D scanner, which is a type of Terrestrial Laser Scanner (TLS) using an articulated 2D laser to gauge and record the distance to objects in the environment. It can capture surface samples over a 360 by 320 degrees’ field of view systematically and a capture typically involves 30-40 million data points with a positional accuracy of +-2 mm stated by the manufacturer. Color information is also collected over the same field by an RGB sensor. The environment that was captured and used in this study was Chalmers Smart Industry laboratory (CSI-lab) that has a 5G network implemented with two radio dots and one radio base station. The two radio dots are installed at a height of 2.9 meters from the floor on a cable ladder. The data collection of spatial data was performed during a three-hour session and there was no other activity in the location meanwhile. To capture the whole area required ten scans in total and the data for the procedure is summarized in Table 10.

Table 10. 3D imaging capture and processing data

Number of scans	Million data points		Data Size		Time duration for	
	Per scan	In total	Raw files	Processed	Capture	Processing
10	10*10^6	100*10^6	650 MB	1,1 GB	3h	3h

The environment in the laboratory mainly contains equipment for teaching, but also some industrial equipment such as a CNC machine. The area is 14.5 by 18 meters (~261 m2) and the height of the ceiling varies between 3 and 4.6 meters. Steel is present to a large extent in the laboratory (pillars, ventilation system, whiteboard, beams that go 40 cm down from the ceiling and CNC machine) as well as sheet metal (cable ladders, fire extinguisher, gates, and interior). The ceiling contains sheet metal under a sound absorbing material which covers 75% of the ceiling, the walls have a first layer in plaster and studs in steel can be found under with a 60 cm spacing and the floor is in concrete with a plastic layer over it.

The Autodesk ReCap software was used to processes the scan data and an example of how it is visualized in the software can be seen in Figure 10 at the second step. The processed data set was used to generate a combined point cloud file covering the entire production area. The combined point cloud was further exported to a neutral data format, .e57 (ASTM) and sub-sequentially imported into Autodesk ReCap for visualization and saved as .rcp data files to be used in Autodesk CAD as blueprints for creating simple CAD representations of the area. An obstacle at this point was how to use the point cloud, i.e., the virtual environment, to perform the ray tracing simulation. It was not possible to use the point cloud as it was since it visualizes the environment with points but the simulation needs solid surfaces to mimic the radio behavior. Therefore, a solution was to build the CAD representation based on the point cloud, step 3 in Figure 10. An automated meshing of the scan data was ruled out since the level of complexity in the machine geometries and because the level of detail was met by the simple geometry sketches. The final result of the spatial data collection can be seen on Figure 10, going from the reality collecting spatial data with 3D laser scanning, creating a point cloud representation and final build a CAD representation based on the point cloud.

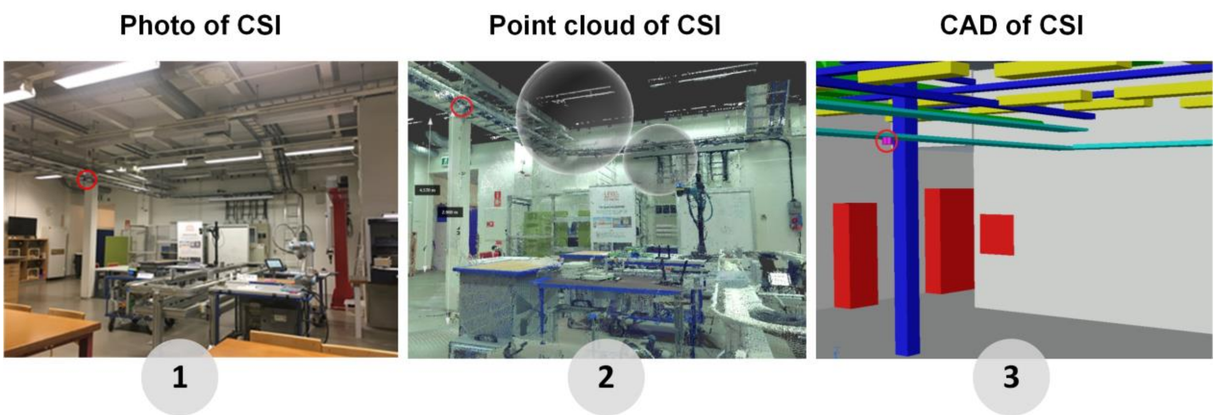

Figure 10. The three steps of capturing the spatial data from reality using 3D laser scanning to a CAD representation (paper 3). The red circle marks the Radio Dot location

4.2.3 Summary

In four out of six appended papers, digital technologies that can make more data available have been used. Technologies applied, the reality, data made available, and the value of data are summarized in Table 11.

Table 11. Digital technology applied, reality, data made available, and value

Digital technologies	Reality	Data	Value
Paper 2			
Communication: Cellular radio network (5G) and sensors. *Computation:* Data analytics (software for big data & advanced algorithms for anomaly detection). *Storage:* Cloud and IoT platform. Data-centre for data storage and distribution. Local client server in the factory for intermediate storage.	A machine with the critical components ball screws, slides, and motor. When they are worn out vibrations, uneven movements and temperature increase can be caused.	Sensor and machine data: • Motor temp. • Torque of motor • Position of slides • Vibration in slides and motors • Tolerance offset • Cycle time	• Real-time data of the machine for the critical components • Data can be collected for longer time-series • Understand the process better • Design vs. reality
Paper 3			
Communication: 3D laser scanning, Faro Focus 3D *Computation:* Point-cloud and CAD software *Storage:* memory cards, scan files, computer	The facility of the workshop where the network should be installed	• Spatial data • Point-cloud data • Location of equipment	• A rapid digital representation of the factory as-is • Placement of equipment in the location
Paper 4			
Communication: 3D laser scanning, Faro Focus 3D *Computation:* Point-cloud and CAD software *Storage:* memory cards, scan files	Workshop with machines and equipment. Eight sub-flows with over 150 individual process steps.	• Spatial data • Point-cloud data with 500 million points • 50 scans	• Gathering data offline about processes in digital form without disrupting production • Capture the production facility as-is • Accurate within less than a centimeter and photorealistic • Supplement other data collection methods • Assistance in model building step for DES
Paper 5			
Communication: Cellular radio network (5G) and sensors. *Computation:* Data analytics (software for big data & advanced algorithms for anomaly detection). *Storage:* Cloud and IoT platform. Data-centre for data storage and distribution. Local client server in the factory for intermediate storage.	A machine with the critical components ball screws, slides, and motor. When they are worn out vibrations, uneven movements and temperature increase can be caused.	Sensor and machine data: • Motor temp. • Torque of motor • Position of slides • Vibration in slides and motors • Tolerance offset • Cycle time	• Real-time data becomes accessible • New ways of existing data • Learn from data how it should be presented • Combination of two expertise domains • Identifying challenges with connecting a machine

4.3 RQ2: FROM DATA TO DECISION SUPPORT

As the data is being made available by the digital technologies, the data can be processed into information for providing decision support for the organization. The same cases presented in RQ1 will now be explained from the perspective where data is made into decision support and what values that bring to the organization.

4.3.1 Connected Machine

From the connected machine, real-time data was sent back to MOST in a smartphone. The system feedback is provided by a mobile operator support system that is already to some extent used in this workshop. It supports the publish-subscribe protocol Message Queuing Telemetry Transport (MQTT) (Weichhart and Wachholder, 2014) which has become popular for many IoT applications. In this support system, real-time data can be presented to the operator to provide information about machine status, but also for presenting trend analysis for a maintenance case. Three functionalities were developed for MOST that were explained in section 4.2.1. Each of these was evaluated by raising the question of how it can support the role of the operator today (question 1 in Table 12) and in the future (question 2 in Table 12). From the focus group interviews, the following results were identified for the requirements of data to support to drive decision-making in the manufacturing context.

Table 12. Results from the evaluation of tolerance offset, vibration data, and cycle times by domain experts from paper 2

Functionality	Question 1: Support today	Question 2: Support in the future
Tolerance offset	It will support flexibility and mobility for the operator. The operator will be informed and can act regardless of location in the workshop.	In the future, it will be even more important for the operator to be able to monitor the machine and preferably anywhere in the workshop.
Vibration data	The value would be to have less unplanned stop because of machining errors. It can also be a valuable input to understand the lifespan of a tool.	To be able to plan better and avoid unplanned stops.
Cycle time	To know the actual cycle time is of interest to the managers and production technicians for follow-up. It is a way to understand how the designed system actually performs in reality.	In the future, this will provide a historical basis to support continuous improvements.

The three functionalities have different importance for different roles in the organization and can fulfill different purposes. To receive data about the tolerance offset is something that can inform the operator that there is a tendency of the machine to go outside the set tolerances. If this data can be available for the operator regardless of location in the workshop, the operator can always be informed of what is happening and do the right actions in order to avoid quality issues driving cost for the company. This is viewed as an even more important aspect of the future role of the operator since it is expected that the operator's responsibility will become more diversified where a high level of flexibility is important. Vibrations in the machine provide the organization with information about the condition of the machine. This data is already being measured by the maintenance organization but by more manual efforts compared to the connected machine that is presented here. To have the historical data of the vibration patterns can provide the organization with better information to understand when an unplanned stop and machining errors can occur as the vibration in the machine changes. Another aspect mentioned was to understand the lifespan of tools in the machine since the vibrations in the machine are correlated to this aspect. The third functionality developed in this project with the connected machine was to receive information about the cycle time. This was termed by the interviewees to not be as important for the operators to receive on a daily basis, but more importantly for the managers and production technicians in order to understand how the system is acting in reality compared to the designed system. With historical data collected and analyzed, this can provide the information for doing continuous improvements of the system.

Challenges encountered when building the data value chain for a predictive maintenance case for a

5G connected machine was addressed in another study, paper 5. Domain experts from manufacturing, information technology, and data analytics were collaborating during this project, and it could be seen that the knowledge gap impacted on the ability to succeed with the data value chain.

4.3.2 Virtual Representation of the Factory

In RQ1, it was explained how data can be enabled of the spatial geometry of the factory to support in layout planning and factory radio design. In this section, it will be lined out and be answered how the data enabled to provide more information for the organization to support decisions.

As the spatial data of the workshop was enabled by 3D laser scanning and a point cloud was created in paper 4, the purpose was to use the virtual environment as a reference for the simulation model. The result of the point cloud provides a photorealistic representation of the environment with an accuracy within less than a centimeter. It was evaluated during and after the simulation project how the virtual representation support in the different steps of a simulation project; in model building and performance, verification and validation, and implementation. A number of benefits but also potential drawbacks were recognized in the different steps of the simulation project summarized in Table 13.

Table 13. Benefits and drawbacks identified for the usage of virtual representation in the different steps of a DES project from paper 4

Project step	Identified potential drawbacks	Benefits of using 3D laser scanning data
Model building and performance	Has a negative effect on model performance due to the large amounts of data being input, which could affect total lead-time.	3D laser scanning helps in pathing and positioning, getting the proper distance relations between functions and components. It could reduce the total amount of iterations required to achieve a validated model, as the accuracy can be improved.
Verification and validation	The realistic visualization could reduce skepticism toward the model, possibly allowing a sub-par model to be validated due to appearing to be more accurate than it is.	3D laser scanning allows the involvement of any personnel as the model is realistic and can be run in real-time, making the model easy to relate to and understand.
Implementation		3D laser scanning can help ensure that changes will fit and work as planned in the real production facility, as well as help reduce resistance to change.

In the model building and performance phase, the point cloud supported in focusing on the model logic rather than creating physical parts, e.g., walls and work benches but also in ensuring that the path of the operator was built in such a way that collisions were avoided. Another aspect was that the accuracy of the model could be reached by fewer numbers of iterations for achieving a valid model. This is on the expense that a large amount of data generated from the point cloud needs to be handled that can impact on the total lead-time. When the model runs it was also easy for the operators working in the real-world system to confirm or add necessary corrections to the path model, supporting in the verification and validation phase of the project. The involvement of personnel for verifying and validating the model can have the potential benefit of less resistance from the organization for the new implementation or change as well as making sure that a solution is developed that is realizable in the production system. Even though the realistic visualization has a lot of identified benefits, a realistic representation may decrease the skepticism from the people being receivers of the model which can hamper the validation and verification of it. For the final step when implementing changes to the system, the realistic visualization enabled by 3D scanning ensures that the designed system will have the same input and output and can also ease the resistance to change among personnel since it provides decision support that is well thought through before implementation. To have the digital version of the production system can also in the future be used together with other digital technologies for planning the layout and changes to the production system, e.g., Virtual Reality, supporting in visualizing the virtual environment.

As was explained in RQ1, the three first step of the offline tool for planning the factory radio design (paper 3) was to create the virtual representation of the factory by capturing the spatial data. When the environment was represented in a 3D version, this could be used for the ray tracing, i.e., physical optics simulation in the environment. A hybrid PO method that has been developed for multi-scale reflector antenna systems (Iupikov, 2014) was used. To enable a time-efficient analysis of the indoor environment of interest for this study, some simplifications of the simulation had to be made: (1) only flat and slowly-curved surfaces of the scatterers are supported (e.g., walls, floor, and ceiling of the room). This can be easily extended by using other methods (MoM, CBFM, MLFMM, etc.) to compute current on electrically small or highly curved objects (Iupikov, 2014), (2) the materials of the walls, floor, and ceiling of the rooms are assumed, perfect electric conductors, (3) mutual coupling effects due to scattering of radio waves in between the walls are negligible, and (4) the radiation pattern of the source is assumed to be a Gaussian pattern (linear-polarized) with the taper -1 dB at θ = 90°. The PO method is implemented in MatLab and to gain more information on the method, the reader is referred to (Iupikov, 2014) and (Iupikov, 2017). Besides the simplifications made for the simulation, the settings of the simulation are documented in Table 14.

Table 14. Simulation parameters setting from paper 3

Simulation parameter	**Setting**
Frequency setting of interest	2.365 GHz
Size of the EM-problem	~332000 mesh cells with a size of ~42×42 mm (~0.33×0.33 wavelength)
No. of radio sources (radio dots)	Two dots at different positions on the ceiling, operating either one by one or simultaneously.
EM modle	The environment consists of 10 walls.
E-field power distribution	Calculated on the measurement path data to enable comparison.

In the simulation, three different cases were explored; (1) both radio dots active, (2) one dot active, and lastly (3) the other radio dot active. To evaluate the results of the simulation and thereby to recognize if the offline tool can automate the planning of factory radio design, measurements of the real network in the CSI-lab was performed with TEMS pocket LTE mobile phone operating in the B40 band to compare the results of the simulations. This was in the same way also performed for the three cases that were simulated. The results of the simulation compared to the path of the measurement are presented in Figure 11 a-b.

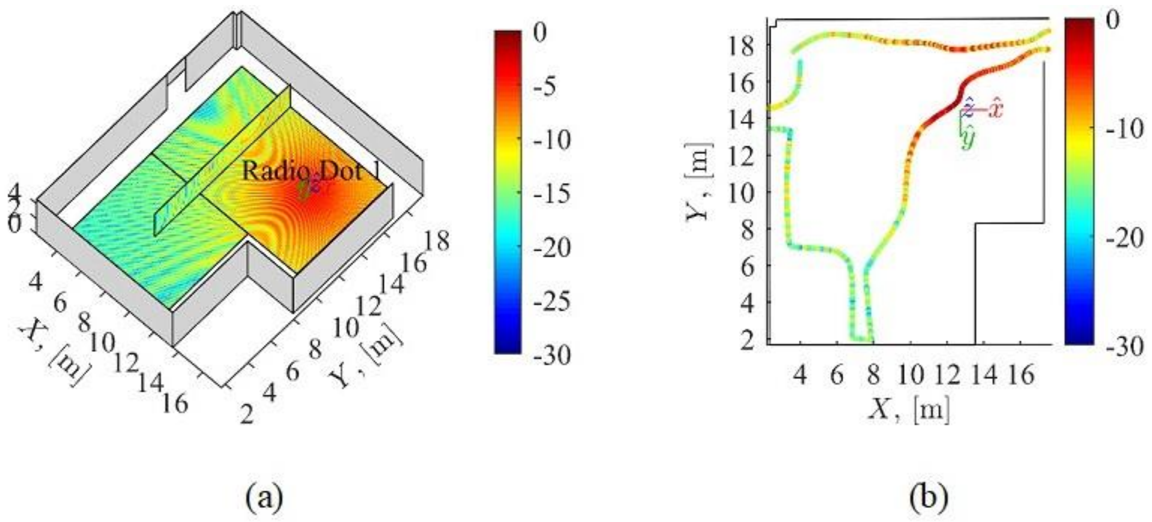

Figure 11. The magnitude of the E-field: (a) on a plane elevated 1 m over the floor, and on a vertical YZ-plane; and (b) along the measurement path in [dB]

For details on the measurements and the comparison between simulations and measurements, the reader is referred to the appended paper 3, "*Digital Twin for Factory Radio Space Design of a 5G Network*". From the results, it could be seen that the PO simulation model mimic the measurement results with reasonable accuracy. The rms error was in the order of 3.2 - 4.6 dB that can be regarded as a good result (De La Roche et al. 2013).

4.3.3 Summary

To answer what additional value that can be gained by decision-support based on data are summarized in Table 15 with data, decision-support, decision, and value.

Table 15. Decision support based on data, decision to support, and values

Data	Decision-support	Decision	Value
Paper 2			
Sensor and machine data: • Motor temp. • The torque of the motor • Position of slides • Vibration in slides and motors • Tolerance offset • Cycle time	*Content:* Value of a data parameter or data over a time-series *Carrier:* Smartphone with a mobile operator support system	*Operation:* required actions and compensation in machine	• Mobility and flexibility of operator with the mobile device • Maintaining quality of product and avoiding scrap
		Maintenance: trend analysis and root cause analysis (RCA)	• Avoid unplanned stops • Possibility to do/plan proactive actions • Be more informed
Paper 3			
• Spatial data • Point-cloud data • Location of equipment	*Content:* Point cloud and CAD-representation of the workshop *Carrier:* General-purpose computer	*Design:* network plan for installation of hardware to ensure radio coverage	• Automated process • Offline planning in a virtual environment • A reliable and available network with low latency
Paper 4			
• Spatial data • Point-cloud data with 500 million points • 50 scans	*Content:* Point cloud and CAD-representation of the workshop *Carrier:* General-purpose computer	*Design:* Change of layout for a more flow oriented production and identify the required investments	• Ensure that the solution can actually be implemented • Reduce resistance to change, as realistic visuals could help to visualize the planned changes • Allows offline work • Reduce the total amount of iterations required to achieve a validated model, as the accuracy can be improved • Involve any personnel as the model is realistic and can be run in real-time
Paper 5			
Sensor and machine data: • Motor temp. • Torque of motor • Position of slides • Vibration in slides and motors • Tolerance offset • Cycle time	*Content:* Anomaly detection in the vibration of sensor data *Carrier:* Smartphone with a mobile operator support system	*Maintenance:* predictive maintenance from data analytics	• Detect anomalies in vibration data in combination with other data parameters. • Graphically display the value-combination as a “tree” (Decision Tree Algorithm), important to increase the understanding of the results of the machine learning algorithms and how data relate to each other

5

ANSWERING RQs

RQ 1) What is the value of collecting data from a production system using digital technologies?

The digital data representing a production system facilitates the ability to know and have an understanding of what is happening in the real-world system and to have the virtual representation of the reality. Two main technologies that have been used are 5G telecommunication, with data managing and analyzing capabilities and 3D laser scanning. For the connected machine, 5G was the enabler for extracting digital data about the machine. This involved data about vibration, motor temperature, cycle time and similar. This gives an understanding of the condition and behavior of the machine in real-time, but also an understanding as the aggregated data is analyzed. The analyzed data over a longer time-series gives values to multiple roles in a production organization. For instance, production technicians designing the system and the maintenance engineers who care for the availability and uptime of the equipment can compare the design of the system to the real system. For the spatial data, it is about providing a digital representation as-is of the system and this can be done with high accuracy. This rapidly provides an understanding of the physical layout and placement of equipment for that environment in an offline mode.

RQ 2) What is the value of decision support based on an increasing amount of digital data from a production system?

The decision support that can be based on more data supports the organization to have a better understanding of what is happening in the real world-system, the trends, and correlation to other data parameters, and enables prediction of the future and how the production system is changing. For the connected machine, more data can support mobility and flexibility of the organization, since they could receive data for decision support regardless of location in the workshop. Proactive actions can then be taken to ensure higher quality yields and better availability of the production equipment. The aggregated data can also enable trend analysis of data to identify previously unknown correlations. The spatial data building the virtual representation of the workshop could support an automated network planning process and support building a simulation model. It also reduces the resistance to change for an organization, since they can be involved in the final solution before implementation.

6

DISCUSSION

The discussion will cover the following topics; production system and current practice of data, the digital technologies and how the applicability of them, an increasing amount of data, data to information, and decision support. This is in relation to what has been important concepts and areas for this thesis. Further on will the quality in the research performed and the stakeholders of this work be discussed.

5.1 THE PRODUCTION SYSTEM AND CURRENT PRACTICE OF DATA

The manufacturing sector is phasing a fourth industrial revolution that will imply a transition towards a bigger focus on data, integration of network-wide ICT, and a higher level of automation with the human still involved in the system (Thoben et al., 2017). Data will come from various sources in the production system and it will be decisive to handle it in real-time for decision-making (Rüßmann et al., 2015). To understand the current system and what the practice of data look like, this section will discuss the availability of data, issues regarding the quality of data, and how the production organization anticipates using data in the future.

In the first study presented in this thesis, paper 1, data were collected with the intended purpose to be used for a simulation model. The approach succeeded in collecting data, but an important last step of the approach was also to show the availability of data needed for representing a production system. Based on the result in Table 4, data about quality yield, tied up capital and required space for storing work in progress were missing. These results motivated a next study to go further into this topic to understand their practice of data today and if they see a need to become more data-oriented. From the interview study with case company A and B in paper 6, the results show that both case companies experienced challenges with data acquisition, but rather of being independent challenges acting in silos the result provide insights that there is an interplay between them. The interplay and how the different challenges impacted each other were especially present at case company A, that has a lack in availability of data as was discussed. Not having procedures of how to handle data in a standardized way, not having systems that are interoperable with each other, and not having a system that is easy to maintain impact on how data is regarded as a reliable source and thereby impacting on the practice of data as decision support to personnel in a production system. The availability of data in the current production system was scarce and thereby impacted on the practice of using data as decision support. From the evaluation performed in paper 2 addressing the importance of big data, it was stated by personnel from the manufacturing domain that most decisions are driven by the intuition and experience by personnel rather than by data that can describe the machine status. The same evaluation expressed that some areas, e.g., maintenance, have started to collect more data to provide trend analysis to support more proactive actions, but at the same time the variety of data available was limited.

The potential values of having more data available were stressed by the participants themselves in all projects and studies, but at the same time, the awareness was high that the data should provide value to the organization to motivate the costs involved in making more data available. Simply having more data may not make the decision-making better according to the respondents' answer and it is crucial to receive the right data in the right form. These statements strengthen the crucial aspect to demonstrate how digital technologies can be used and how the data provides benefits to the organization. Since the theory presented in this thesis and examples from other sectors show that having simply more data can provide value even though the use of it is not obvious in advance. As both theory and practice have demonstrated, there is a need to identify the value of data for decision-making, which is the stated aim of this thesis. By answering the two research questions, that cover separate parts of the data value chain, the thesis can get closer to the aim.

5.2 DIGITAL TECHNOLOGIES

The digital technologies have had and still have a tremendous impact on society and industries. As presented in Figure 5 from Hilbert and López (2011), the digital and analog technologies are divided into communication, computation, and storage. In this work, an important aspect has been the digital technologies that can make the data generated at different levels in the production system available to the organization for supporting decision-making. This section will go further into the topic and discuss the digital technologies that have been used, i.e., 5G telecommunication and 3D laser scanning and how they support in digitalizing data about the production system.

By deploying 5G it was possible to extract both data that was already stored in the machine computer, collected for other purposes, and by adding external sensors additional data about vibration and temperature could also be collected. To connect sensors and/or devices is supported by the flexibility and mobility that 5G provides to the shop floor. The data in itself, so-called raw data, can provide the benefits to the organization to know the status of the machine, for instance, what the vibration looks like or that something has caused a stop in the machine.

3D laser scanning captures and digitalizes spatial data of the production system and two of the studies utilized this technology for making data available. In both studies, the main objective was to capture spatial data of the production system, but they differ to some degree depending on the end purpose. In paper 3, the intention was to support the design of a network for an indoor environment, and here the information about the facility, placement of equipment but also the material present in the environment are important factors since they impact on the radio coverage for an indoor environment. In paper 4, it was mainly the data that can describe the facility and where equipment is placed that is important in order to support building a simulation model representing the production system. 3D laser scanning can support in any project or study that requires an accurate understanding of the existing facility that is going to be evaluated or changed since it is a rapid method for generating the virtual representation. The spatial data is providing a common basis for what the environment looks like in reality represented in a virtual version that is not interfering with the ongoing production.

5.3 INCREASING AMOUNT OF DATA

Increasing amount data and big data has been a topic throughout this thesis, and referring to the definition it involves data in big volumes, fast in retrieval and processing, and collected from a big variety of resources (Laney, 2001). Besides that, other aspects that have been presented as important for the big data era are value, variability, and veracity (Gandomi and Haider, 2015). In the new era where the technologies can support the new type of data, there is an opportunity to collect as much data as possible, since the secondary use of data has become more important. As Mayer-Schönberger and Cukier

(2013) put it, much of the value in data comes from the secondary use of it and if that can be considered for at an early stage, this motivates the collection of multiple data streams and many data points in each stream. This is because the cost does not increase by significant numbers when the volume of data increases and can therefore further motivate the collection of data.

The studies presented in the results chapter strengthen the idea that decision-making driven by data are better decisions and more data made available can support the organization. In the case of the connected machine, the technology even enabled new data sources that have not been available in digital form before to the production organization. It was also seen in the current state, that there is a lack of available data that are free from quality issues to support the decision-making. The results tell that the current practice of data in the manufacturing sector is scarce. The understanding, however, was that more data and data with high quality can support the organization to do more informed decisions. As one of the interviewees puts it "*data acquisition is very important for this initiative, it is a no-brainer.*"

In the studies where digital technologies have been used, it can be seen that they have collected and enabled more data representing the reality of the production system. In the case of the connected machine, several data streams were collected and combined to gain insights about correlation supporting numerous roles in the production organization depending on use. For the studies where the spatial data was collected, the procedure is the same regardless of the intended use and can be used for multiple purposes, which indicates that data have a secondary use.

5.4 DATA TO INFORMATION

The big data era is not only about having more data but also to make meaningful insights from the data. With the new technologies and ways of handling data, it will be possible to process and analyze it for the further use of it. They will support for the analysis of a wide range of different data types, such as textual, video and audio content. If you go even further to using the machine learning application, it can create statistical models that fit, optimize and predict the data in a rapid matter (Davenport, 2014). Some scholars stress that data itself not necessarily give value to the organization, but it is when it has been transformed to information and provide insights as the real value is achieved (Chatfield et al. 2015; Li et al. 2014).

In the case of the connected machine, several data streams were collected. The data streams were selected since they were termed to be of interest to support numerous different roles that are related to the production organization. The various data streams collecting different physical signals, such as the vibration or temperature, were combined at the analytic phase of the study. What was strived for here was to identify new correlations between different data streams that have not been known by the organization before. An example of a question raised was "can the combination of vibration and temperature data tell the organization that the machine condition is as desired, or is a trend in the data indicating the need of maintenance or other actions?". This mindset and approach are much aligned to what the big data era will imply, to combine new data streams containing big volumes of data points to identify new correlations to gain new insights about processes and operations. The virtual representation is another example of where the data itself, depending on need, can be used multiple times for gaining more insights. In this thesis, the same procedure for collecting the data has been done, but the processing and change of the virtual version have been adapted to the two purposes when building a simulation model and planning a network.

5.5 DECISION SUPPORT

The fourth revolution is taking place in the manufacturing sector and the role of the operator in this system is anticipated to change. The role of the humans is expected to shift from operating a single machine towards a role and responsibility for making strategic decisions to manage the system and also be a problem solver that can cope with the changing circumstances (Hermann et al., 2016). As was stated by Koŝcielniak and Puto (2015), decision-making involves making a selection in a situation that otherwise would not be resolved. In this context, the data is providing support building the basis to make an informed decision. The results in this thesis strengthen the idea that data can provide decision support in the manufacturing environment to make more informed decisions. As the two authors McAfee and Brynjolfsson (2012) state *"data-driven decisions are better decisions – it's as simple as that"* there is much in this saying to learn by the manufacturing sector. There is potential as the technologies become more mature and the organizations can see the impact and value of having more data available to understand their processes and operations. The intuition will be complemented by data-driven decision support to make informed decisions.

Information presented to a domain expert should involve facts, concepts, and theories that are relevant to the domain of the problem (Byström and Järvelin, 1995). In paper 5, the topic of how the gap between two separate domain expert groups needed to be bridged in order to build the data value chain was addressed. According to the literature about the knowledge gap between domain experts in different fields, the study could confirm that this gap exists. Based on the experiences from the study, the authors suggest the following improvements in order to ease the creation of a data value chain for future cases: (1) agile work cycle, (2) know what parts that are self-comparable and what parts that are not relevant, (3) connect the data to relevant metadata depending on products, components, machines, batches, or product families, and (4) experiment with data but let the manufacturing process experts guide these experiments. To make the decision support relevant to the domain expert should be improved with the efforts of point 2, 3, and 4. It was seen from the study that the domain knowledge needs to be part of the development of decision support in order to evaluate if it is providing any new and relevant information. An example from the study concerns the resolution and frequency data is presented in. If the operations are in second, the frequency of presenting data in intervals of minutes is too big and will not tell the organization when something unexpected is happening. For this new area where IT related and manufacturing expertise are going to be combined, it will be crucial to work together and address the four points presented here for creating the data value chain.

From the studies, it was identified that the additional values provided from the data involve the ability to know more about when something happens, provide better insights for planning, but also in the communication with external partners. Another aspect is the democratization of who can impact on the decision made since everyone is provided with the same decision support. It also supports in the design of the system, both seen in the planning of a network, building the simulation model, and for the evaluation of how well the real-system corresponds to the designed system. This brings the value that interference with the real-system could be reduced. Values mentioned was the enabling of higher flexibility of personnel when data of machine status can be presented anywhere at any time. In paper 2 it was stressed, however, that it is vital for the data made available is the right data with the right resolution. This statement strengthens what was explained in section 2.6 that information should represent the event that will occur and how it should be handled. It involves the structure, properties, and requirements of the problem (Byström and Järvelin, 1995) and information should increase the receivers understanding of the problem.

5.6 QUALITY IN RESEARCH

To state the current practice of data, was an outcome of an effort with another intended purpose and in a next stage followed up with more in-depth interviews. The findings have been supported by the research literature strengthen the generalizability. A limitation of the studies performed is, however, the scope and scale of the study. To state, the current practice has included three companies and most results have been based on the employees' experience and understanding. For the focus interview in paper 2, a selection of people with different responsibilities and purpose of using data were selected. To make the results even more solid would require multiple focus interviews involving a bigger variety of roles from the company.

To answer RQ 1 and 2, a change was made to the system introducing and applying digital technologies to make more data available. To evaluate and identify how this provide value to the organization, qualitative data collection methods such as interviews, observations, and also participation in the study were used. Even stronger statements would be to have quantitative measures of how these new data could provide value, for example increasing productivity, but it was termed that it was more important to identify the values stated by the production organization at this point.

Most studies have been performed by a group of researcher, especially the technical implementation in project B that has also involved more roles from IT/telecommunication and the manufacturing field. Interview guides have been developed based on topics identified in the research field and that has been termed to be of interest to investigate further. Being more than one researcher while interviewing, ensured more than one perspective to the study. In paper 6, semi-structured interviews were used and the researcher did the first analysis by themselves according to a number of areas that were decided on beforehand. This was in a later stage discussed together and the intention was to not impact each other in the analysis phase to make the findings more trustworthy.

The papers appended to this thesis have all went through a peer review process, where at least two researchers have reviewed them. This process has provided very valuable comments about the content, how the studies were performed and the conclusions derived. The comments have revealed where the statements have not been clear enough and where the expectations received from the introduction was not aligned with the methodology and results presented. By highlighting where expected and actual content deviates, the quality of the papers could be increased. Most of the papers appended are conference proceedings, which have provided the opportunity to present the paper to other scholars in the field as well. The discussions on these occasions have provided a good opportunity to receive insights and comments about the research performed and to understand how it can contribute to the research community. Paper 3 has been submitted to a journal.

To ensure the credibility of the research, the transcribed and documented interview answers have always been shared with the interviewees for them to review. For each of the paper appended, the industry partners have been co-authoring and thereby also confirmed the content that is presented. For the transferability, it was already mentioned that it can be questioned because of the limited number of interviews and companies studied. However, it was also stated that the findings are aligned with the current discussion in the research field that can strengthen the results presented. Another aspect of the transferability is the technical implementation and change to the system that was made with the intended purpose to be scalable and be possible to be performed twice. This is a reason why these papers put much emphasis on describing the process and how the technical implementation was performed to make it transferable to similar cases. In the same way, the intended purpose to reduce the dependability of the studies was to make it possible to observe the same phenomena twice. For the confirmability, the bias has been reduced by performing the studies by more than one research and the review process have also made the content less bias.

5.7 STAKEHOLDERS

The intended stakeholder and receiver of this work is the manufacturing domain. Both within the research community and the practitioner in industry. The purpose has been to state how an increasing amount of data can provide value to the production organization. As the title suggests, the value of data is increased by the results that have been presented here. Firstly, from the results stating the current practice of data that shows on an understanding of why data to support decisions is important and what the consequences are when data is not available. Secondly, in the studies where the delta before and after using digital technologies has been described and the values have been identified from having an increased amount of data. So the intended receiver of this work is someone that seeks to understand how the value of the data in a production system can be used in order to gain value from it.

7

CONCLUSIONS

Digital technologies are providing society and industry with new opportunities for making data available in digital form. Amongst industries, the manufacturing domain is one of them and in order to make them adopt these technologies it will be crucial to show how the data provide values. From the studies that have been exploring the current practice of data to represent the production system it was seen that the availability of data was scarce, and in the cases where data existed, the acceptance of the data was low or questioned. The current practice of data use in the production system and the awareness that decision-making should be based on more data motivated the research questions. The aim of this thesis was therefore to identify the value of data for decision-making.

The first question in this thesis concerns how more digital data being available can provide value to the manufacturing organization. The values identified for the connected machine are the status for the critical machine components, comparison between the behavior of the real-world system and the designed system, understanding the processes better, and analysis of aggregated data over a longer time series to learn from data but also to enrich the understanding of what the challenges are when connecting a machine. For the spatial data collection, the technology captures a production system as-is and it is very accurate and photorealistic. It can also supplement other data collection methods and assist in the next step when it is intended to make a change in a production system.

"To measure more is to know more" (McAffee and Brynjolfsson, 2012) is a statement that has been a cornerstone throughout this thesis and has also been justified by the results presented to answer research question 2. In the case of the connected grinding machine, more data about the condition of the machine can support the operator in daily operations, the maintenance organization in making better predictions of the required actions, and can provide the production engineer with a feedback loop of how the system is acting in reality compared to the intended behavior. In the case of providing more spatial data about an environment that can be presented in a virtual representation, the data has supported the involvement of several different roles within the company and allowed the building of a realistic simulation model. When designing the network, data support can provide the spatial information in an offline mode that can make the factory radio design process automated instead of being performed in an ad hoc manner based on experience and intuition. It also allows the planning of changes in an offline environment that will reduce disturbance of production.

REFERENCES

Adam, E. E., Jr., & Swamidass, P. M. (1989). Assessing Operations Management from a Strategic Perspective. *Journal of Management, 15*(2), 181-203. doi:10.1177/014920638901500204

Akyildiz, I. F., Melodia, T., & Chowdhury, K. R. (2007). A survey on wireless multimedia sensor networks. *Computer Networks, 51*(4), 921-960. doi:10.1016/j.comnet.2006.10.002

ASTM. (2011). ASTM E2807-11 Standard Specification for 3D Imaging Data Exchange. Version 1.0, ASTM International, West Conshohocken, PA, 2011, doi:10.1520/E2807-11

Baillieul, J., & Antsaklis, P. J. (2007). Control and communication challenges in networked real-time systems. *Proceedings of The IEEE, 95*(1), 9-28. doi:10.1109/JPROC.2006.887290

Behrendt, A., Kadocsa, A., Kelly, R., & Schirmers, L. (2017). How to achieve and sustain the impact of digital manufacturing at scale. *McKinsey&Company Quartley*. Retrieved from https://www.mckinsey.com/business-functions/operations/our-insights/how-to-achieve-and-sustain-the-impact-of-digital-manufacturing-at-scale

Bellgran, M. (1998). *Systematic design of assembly systems -Preconditions and design process planning.* (PhD), Linköping University

Bellgran, M., & Säfsten, K. (2005). *Produktionsutveckling*. Lund, Sweden: Studentlitteratur.

Bennett, B., Boddy, M., Doyle, F., Jamshidi, M., & Ogunnaike, T. (2004). Assessment Study on Sensors and Automation in the Industries of the Future. Reports on Industrial Controls, Information Processing, Automation, and Robotics. doi:10.2172/1218800

Bennett, D. (1986). *Production systems design*. London, UK: Butterworth-Heinemann.

Berglund, J., Lindskog, E., & Johansson, B. (2016). *On the trade-off between data density and data capture duration in 3D laser scanning for production system engineering.* Paper presented at the Procedia CIRP.

Brous, P., Janssen, M., Schraven, D., Spiegeler, J., & Duzgun, B. C. (2017). *Factors influencing adoption of IoT for data-driven decision making in asset management organizations.* Paper presented at the 2nd International Conference on Internet of Things, Big Data and Security (IoTBDS 2017), Porto, Portugal

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. New York, NY, USA: WW Norton & Company.

Bughin, J., Chui, M., & Manyika, J. (2015). *An executive's guide to the Internet of Things*. Retrieved from https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/an-executives-guide-to-the-internet-of-things:

Buhl, H. U., Röglinger, M., Moser, F., & Heidemann, J. (2013). Big Data. *Business & Information Systems Engineering, 5*(2), 65-69. doi:10.1007/s12599-013-0249-5

Burrell, G., & Morgan, G. (1979). *Sociological paradigms and organisational analysis*. London: Heinemann.

Byström, K. (1999). *Task complexity, information types and information sources: examination of relationships.* (PhD), Tampere University,

Byström, K., & Järvelin, K. (1995). Task complexity affects information seeking and use. *Information processing & management, 31*(2), 191-213. doi:10.1016/0306-4573(95)80035-R

Bärring, M., Berglund, J., Johansson, B., Stahre, J., Iupikov, O., Glazunov, A. A., . . . Friis, M. (2019). Offline Planning Tool for Factory Radio Space Design of a 5G Network. *Submitted to Journal*.

Bärring, M., Flores-García, E., Johansson, B., Bruch, J., & Wahlström, M. (2018). *Challenges of Data Acquisition for Simulation Models of Production Systems in Need of Standards.* Paper presented at the Procedia of 2018 Winter Simulation Conference (WSC), Gothenburg, Sweden.

Bärring, M., Lundgren, C., Åkerman, M., Johansson, B., Stahre, J., Engström, U., & Friis, M. (2018). *5G Enabled Manufacturing Evaluation for Data-Driven Decision-Making.* Paper presented at the Procedia CIRP.

Bärring, M., Nåfors, D., Henriksen, D., Olsson, D., Johansson, B., & Larsson, U. (2017, 3-6 Dec. 2017). *A VSM approach to support data collection for a simulation model.* Paper presented at the 2017 Winter Simulation Conference (WSC).

Chapanis, A. (1996). *Human factors in systems engineering*. New York, NY, USA: John Wiley.

Chatfield, A., Reddick, C., & Al-Zubaidi, W. (2015). *Capability challenges in transforming government through open and big data: Tales of two cities.* Paper presented at the 36th International Conference on Information Systems, Fort Worth, TX, USA.

Chen, M., Mao, S., & Liu, Y. (2014). Big Data: A Survey. *Mobile Networks and Applications, 19*(2), 171-209. doi:10.1007/s11036-013-0489-0

Churchman, C. W. (1968). *The systems approach* (Vol. 231). New York, NY, USA: Dell.

CIRP. (1990). Nomenclature and definitions for manufacturing systems (English Language Version). *Annals of the CIRP, 39*(2), 735-742.

Cortada, J. W. (2004). *The digital hand: How computers changed the work of American manufacturing, transportation, and retail industries*. New York, NY, USA: Oxford University Press.

Creswell, J. W. (2014). *Research design: Qualitative and Quantitative Approach* (4th ed.). Thousand Oak, USA: SAGE Publications.

Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. London: Sage.

Davenport, T. H. (2006). Competing on analytics. *Harvard Business Review, 84*(1), 98-107.

Davenport, T. H. (2014). *Big data @ work: dispelling the myths, uncovering the opportunities*. Boston, Massachusetts, USA: Harvard Business Review Press.

Davenport, T. H., & Prusak, L. (1998). *Working knowledge: How organizations manage what they know*. Boston, Massachusetts, USA: Harvard Business Press.

Davis, J., Edgar, T., Porter, J., Bernaden, J., & Sarli, M. (2012). Smart manufacturing, manufacturing intelligence and demand-dynamic performance. *Computers & Chemical Engineering, 47*, 145-156. doi:10.1016/j.compchemeng.2012.06.037

De La Roche, G., Alayón-Glazunov, A., & Allen, B. (2013). *LTE-advanced and next generation wireless networks: channel modelling and propagation*. Chichester, U.K.: John Wiley & Sons.

Ericsson AB. (2011). *More than 50 billion connected devices*. In Ericsson White Paper. Retrieved from https://www.ericsson.com:

Ericsson AB. (2017). *5G Systems: Enabling the transformation of industry and society*. In Ericsson White Paper. Retrieved from https://www.ericsson.com:

Esmaeilian, B., Behdad, S., & Wang, B. (2016). The evolution and future of manufacturing: A review. *Journal of Manufacturing Systems, 39*, 79-100. doi:10.1016/j.jmsy.2016.03.001

Evans, J. R., & Lindsay, W. (1989). *The Management and Quality Control*. St. Paul: West Publishing Company.

Frezzini, F. R., Sachan, R., & Azimi, M. (2011). *Systems engineering tools and methods*. Boca Raton: CRC Press.

Fässberg, T., Fasth, Å., & Stahre, J. (2012). *A classification of carrier and content of information.* Paper presented at the 4th CIRP Conference on Assembly Technologies and Systems (CATS 2012), Ann Arbor, Michigan, USA.

Galliers, R. (1987). *Information analysis: selected readings*. Sydney, Australia: Addison-Wesley.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137-144. doi:10.1016/j.ijinfomgt.2014.10.007

Ghauri, P. N., & Grønhaug, K. (2010). *Research methods in business studies: A practical guide* (4th ed.). Harlow, England: Pearson Education.

Gillon, K., Brynjolfsson, E., Mithas, S., Griffin, J., & Gupta, M. (2012). *Business analytics: radical shift or incremental change?* Paper presented at the 33rd International Conference on Information Systems, Orlando, FL, USA.

Gorecky, D., Schmitt, M., Loskyll, M., & Zühlke, D. (2014). *Human-machine-interaction in the industry 4.0 era.* Paper presented at the 12th IEEE International Conference on Industrial Informatics (INDIN), Porto Alegre, Brazil

Gungor, V. C., & Hancke, G. P. (2009). Industrial wireless sensor networks: Challenges, design principles, and technical approaches. *IEEE Transactions on Industrial Electronics, 56*(10), 4258-4265. doi:10.1109/TIE.2009.2015754

Hermann, M., Pentek, T., & Otto, B. (2016). *Design principles for industrie 4.0 scenarios.* Paper presented at the 49th Hawaii International Conference on System Sciences (HICSS), Koloa, HI, USA

Hilbert, M. (2012). How much information is there in the "information society"? *Significance, 9*(4), 8-12. doi:10.1111/j.1740-9713.2012.00584.x

Hilbert, M., & Lopez, P. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. *Science, 332*(6025), 60-65. doi:10.1126/science.1200970

Hill, T. (1985). *Manufacturing strategy: The strategic management of the manufacturing function* (2nd ed.). London: Macmillan Press Ltd.

Hounshell, D. (1984). *From the American system to mass production, 1800-1932*. Baltimore: John Hopkins University Press.

Hubka, V., & Eder, W. E. (1988). *Theory of technical systems*. Berlin: Springer-Verlag.

Iupikov, O. A. (2017). *Digital Beamforming Focal Plane Arrays for Radio Astronomy and Space-Borne Passive Remote Sensing.* (PhD), Chalmers University of Technology

Iupikov, O. A., Maaskant, R., Ivashina, M. V., Young, A., & Kildal, P.-S. (2014). Fast and accurate analysis of reflector antennas with phased array feeds including multiple reflections between feed and reflector. *IEEE Transactions on Antennas and Propagation, 62*(7), 3450-3462. doi:10.1109/TAP.2014.2320529

Jonsson, P., & Mattsson, S.-A. (2009). *Manufacturing, planning and control*. New York, NY, USA: McGraw-Hill Education.

Kagermann, H., Helbig, J., Hellinger, A., & Wahlster, W. (2013). *Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Securing the future of German manufacturing industry*. Retrieved from https://docplayer.net/254711-Securing-the-future-of-german-manufacturing-industry-recommendations-for-implementing-the-strategic-initiative-industrie-4-0.html

Kang, H. S., Lee, J. Y., Choi, S., Kim, H., Park, J. H., Son, J. Y., . . . Do Noh, S. (2016). Smart manufacturing: Past research, present findings, and future directions. *International Journal of*

Precision Engineering and Manufacturing-Green Technology, 3(1), 111-128. doi:10.1007/s40684-016-0015-5

Kauffman, D. L. (1980). *Systems one: An introduction to systems thinking*. Minneapolis, MN, USA: Future Systems.

Kelemen, M. L., & Rumens, N. (2008). *An introduction to critical management research*. London: Sage.

Kendrick, J. W. (1986). *Improving company productivity: handbook with case studies*. Baltimore: Johns Hopkins University Press.

Ketokivi, M., & Mantere, S. (2010). Two strategies for inductive reasoning in organizational research. *Academy of Management Review, 35*(2), 315-333. doi:10.5465/amr.35.2.zok315

Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the Acm, 53*(1), 148-152. doi:10.1145/1629175.1629210

Klein, L., Li, N., & Becerik-Gerber, B. (2012). Imaged-based verification of as-built documentation of operational buildings. *Automation in Construction, 21*, 161-171. doi:10.1016/j.autcon.2011.05.023

Koŝcielniak, H., & Puto, A. (2015). *BIG DATA in Decision Making Processes of Enterprises.* Paper presented at the International Conference on Communication, Management and Information Technology (ICCMIT), Prague, Czech Republic.

Kurose, J. F., & Ross, K. W. (2013). *Computer networking: a top-down approach* (6 ed.). Boston, Massachusetts, USA: Pearson Education.

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note, 6*(70), 1.

Li, L., Zhang, C., & Zheng, G. (2014, 7-9 August). *Identifying Important Skill Sets to Support Healthcare Data Processing and Analytics: An Empirical Examination of Perceptions from HIT Practitioners.* Paper presented at the 20th Americas Conference on Information Systems, Savannah, GA, USA.

Lu, Y., Morris, K. C., & Frechette, S. (2015). *Standards landscape and directions for smart manufacturing systems.* Paper presented at the 2015 IEEE International Conference on Automation Science and Engineering (CASE), Gothenburg, Sweden

Luminol. (2018). Retrieved from https://github.com/linkedin/luminol

Löfgren, K.-G. (1983). *Produktion - teknik och ekonomi*. Malmö: LiberHermods.

Machlup, F. (1983). Semantic Quirks in Studies of Information. In F. Machlup and U. Mansfield (Ed.), *The Study of Information* (pp. 641-671). New York, NY, USA: John Wiley.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity*. Retrieved from https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation:

Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data - A Revolution That Will Transform How We Live, Think and Work*. London: John Murray.

McAfee, A., & Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard Business Review, 90*(10), 60-68.

Miller, H. (1996). The Multiple Dimensions of Information Quality. *Information Systems Management, 13*(2), 79-82. doi:10.1080/10580539608906992

Miller, P., & Roth, A. (1988). Manufacturing strategies. Executive summary of the 1988 manufacturing futures survey. In *Manufacturing roundtable research report*: Boston University.

Monostori, L. (2014). Cyber-physical production systems: Roots, expectations and R&D challenges. *Procedia CIRP, 17*, 9-13. doi:10.1016/j.procir.2014.03.115

Moyne, J. R., & Tilbury, D. M. (2007). The emergence of industrial control networks for manufacturing control, diagnostics, and safety data. *Proceedings of The IEEE, 95*(1), 29-47. doi:10.1109/JPROC.2006.887325

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science, 5*(1), 14-37. doi:10.1287/orsc.5.1.14

Nåfors, D., Bärring, M., Estienne, M., Johansson, B., & Wahlström, M. (2018). Supporting Discrete Event Simulation with 3D Laser Scanning and Value Stream Mapping: Benefits and Drawbacks. *Procedia CIRP, 72*, 1536-1541.

Open CV (Open Source Computer Vision Library). (2018). Retrieved from: https://opencv.org/

Osseiran, A., Boccardi, F., Braun, V., Kusume, K., Marsch, P., Maternia, M., . . . Taoka, H. (2014). Scenarios for 5G mobile and wireless communications: the vision of the METIS project. *IEEE Communications Magazine*, *52*(5), 26-35. doi:10.1109/MCOM.2014.6815890

Osseiran, A., Braun, V., Hidekazu, T., Marsch, P., Schotten, H., Tullberg, H., . . . Schellman, M. (2013). *The foundation of the mobile and wireless communications system for 2020 and beyond: Challenges, enablers and technology solutions.* Paper presented at the 2013 IEEE 77th Vehicular Technology Conference (VTC Spring), Dresden, Germany

Patel, M., Shangkuan, J., & Thomas, C. (2017). *What's new with the Internet of Things?* Retrieved from https://www.mckinsey.com/industries/semiconductors/our-insights/whats-new-with-the-internet-of-things: https://www.mckinsey.com/industries/semiconductors/our-insights/whats-new-with-the-internet-of-things

Persson, P., & Angelsmark, O. (2015, June 2-5). *Calvin–merging cloud and IoT.* Paper presented at the 6th International Conference on Ambient Systems, Networks and Technologies (ANT 2015), London, United Kingdom

Poleto, T., de Carvalho, V. D. H., & Costa, A. P. C. S. (2015). *The roles of big data in the decision-support process: an empirical investigation.* Paper presented at the International conference on decision support system technology, Belgrade, Serbia.

Porter, M. E. (1980). *Competitive strategy: Techniques for analyzing industries and competitors*. New York, NY, USA: Free Press.

Pospiech, M., & Felden, C. (2012). *Big data – a state-of-the-art.* Paper presented at the Proceedings of the Eighteenth Americas Conference on Information Systems, Seattle, Washington, USA.

Rajkumar, R., Lee, I., Sha, L., & Stankovic, J. (2010, 13-18 June). *Cyber-physical systems: the next computing revolution.* Paper presented at the 47th ACM/IEEE Design Automation Conference (DAC), Anaheim, CA, USA

Rampersad, H. K. (1994). *Integrated and Simultaneous Design for Robotic Assembly*. Chichester, England: John Wiley & Sons Ltd.

Renwick, J. T., & Babson, P. E. (1985). Vibration Analysis - A Proven Technique as a Predictive Maintenance Tool. *IEEE Transactions on Industry Applications* (2), 324-332. doi:10.1109/TIA.1985.349652

Robertson, N., & Perera, T. (2002). Automated data collection for simulation? *Simulation Practice and Theory, 9*(6-8), 349-364. doi:10.1016/S0928-4869(01)00055-6

Robinson, S. (2004). *Simulation: the practice of model development and use*. Chichester, UK: Wiley.

Robinson, S., & Bhatia, V. (1995). *Secrets of successful simulation projects.* Paper presented at the 27th Conference on Winter Simulation, Arlington, Virginia, USA

Rodrique, J.-P., Farra, F., Jun, N., Ferraz, J. C., & Alcorta, L. (2014). *The Future of Manufacturing: Driving Capabilities, Enabling Investments*. Retrieved from https://gvcc.duke.edu/wp-content/uploads/Future_of_Manufacturing_Driving_Capabilities.pdf

Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., & Harnisch, M. (2015). *Industry 4.0: The future of productivity and growth in manufacturing industries*. Retrieved from https://www.bcg.com/en-nor/publications/2015/engineered_products_project_business_industry_4_future_productivity_growth_manufacturing_industries.aspx: https://www.bcg.com/publications/2015/engineered_products_project_business_industry_4_future_productivity_growth_manufacturing_industries.aspx

Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research methods for business students* (7th ed.). Harlow, England: Pearson education.

Schroeder, R. G., Anderson, J. C., & Cleveland, G. (1986). The content of manufacturing strategy: an empirical study. *Journal of operations management, 6*(3-4), 405-415. doi:10.1016/0272-6963(86)90013-6

Schuh, G., Anderl, R., Gausemeier, J., ten Hompel, M., & Wahlster, W. (2017). *Industrie 4.0 Maturity Index*. Retrieved from https://www.acatech.de/Publikation/industrie-4-0-maturity-index-managing-the-digital-transformation-of-companies/

Shannon, C. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal, 27*, 379-423 & 623-656. doi:10.1002/j.1538-7305.1948.tb01338.x

Shipp, S. S., Gupta, N., Lal, B., Scott, J. A., Weber, C. L., Finnin, M. S., . . . Thomas, S. (2012). *Emerging global trends in advanced manufacturing*. Retrieved from https://apps.dtic.mil/dtic/tr/fulltext/u2/a558616.pdf

Skinner, W. (1978). *Manufacturing in the corporate strategy*. New York, NY, USA: John Wiley & Sons.

Skoogh, A., & Johansson, B. (2008, 7-10 December). *A methodology for input data management in discrete event simulation projects.* Paper presented at the Proceedings of the 40th Conference on Winter Simulation, Miami, FL, USA.

Skoogh, A., Perera, T., & Johansson, B. (2012). Input data management in simulation–Industrial practices and future trends. *Simulation Modelling Practice and Theory, 29*, 181-192. doi:10.1016/j.simpat.2012.07.009

Skyttner, L. (2005). *General systems theory: problems, perspectives, practice*. Singapore: World scientific.

SMLC. (2011). *Implementing 21st century smart manufacturing: Workshop Summary Report*.

Son, Y. K., & Park, C. S. (1987). Economic measure of productivity, quality and flexibility in advanced manufacturing systems. *Journal of Manufacturing Systems, 6*(3), 193-207. doi:10.1016/0278-6125(87)90018-5

Strong, D. M., Lee, Y. W., & Wang, R. Y. (1997). Data quality in context. *Communications of the ACM, 40*(5), 103-110. doi:10.1145/253769.253804

Swamidass, P. M. (1986). Manufacturing strategy: its assessment and practice. *Journal of operations management, 6*(3-4), 471-484. doi:10.1016/0272-6963(86)90018-5

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), 4-19. doi:10.20965/ijat.2017.p0004

Trochim, W., Donnelly, J. P., & Arora, K. (2016). *Research methods: The essential knowledge base*. Boston, Massachusetts, USA: Cengage Learning.

Varady, T., Martin, R. R., & Cox, J. (1997). Reverse engineering of geometric models—an introduction. *Computer-aided design, 29*(4), 255-268. doi:10.1016/S0010-4485(96)00054-1

Weichhart, G., & Wachholder, D. (2014). *On the interoperability contributions of S-BPM.* Paper presented at the International Conference on Subject-Oriented Business Process Management, Eichstätt, Germany

Wheelwright, S. C. (1984). Manufacturing strategy: defining the missing link. *Strategic management journal, 5*(1), 77-91. doi:10.1002/smj.4250050106

Wu, B. (1994). *Manufacturing systems design and analysis* (2nd ed.). London, UK: Chapman and Hall.

Zhou, Z.-H., Chawla, N. V., Jin, Y., & Williams, G. J. (2014). Big data opportunities and challenges: Discussions from data analytics perspectives [discussion forum]. *IEEE Computational Intelligence Magazine, 9*(4), 62-74. doi:10.1109/MCI.2014.2350953

Åkerman, M., Lundgren, C., Bärring, M., Stahre, J., Folkesson, M., Berggren, V., Engström, U., & Friis, M. (2018). *Challenges Building a Data Value Chain to Enable Data-Driven Decisions: A Predictive Maintenance Case in 5G-Enabled Manufacturing.* Paper presented at the Procedia 28th International Conference in Flexible Automation and Intelligent Manufacturing (FAIM), Columbus, OH, USA.