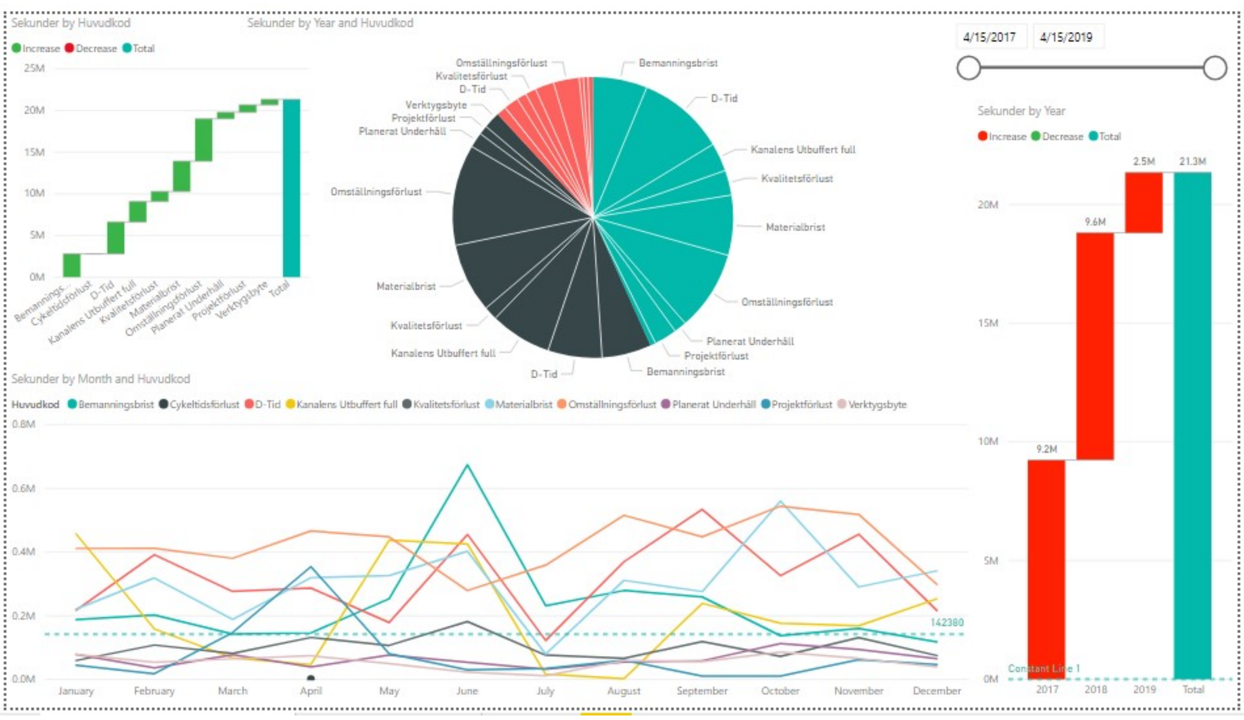
CHALMERS
UNIVERSITY OF TECHNOLOGY



Creation of a Local Data Analytics Center

Functional requirements of existing data analytic tools

Bachelor's Thesis in Mechanical Engineering

CHRISTINA PHAN
EBBA ANDRÉN

Department of Industrial and Materials Science
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2019

BACHELOR'S THESIS 2019

Creation of a Local Data Analytics Center

Functional requirements of existing data analytic tools

CHRISTINA PHAN
EBBA ANDRÉN

Department of Industrial and Materials Science
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2019

Creation of a Local Data Analytics Center
Functional requirements of existing data analytic tools
Bachelor's Thesis in Mechanical Engineering
CHRISTINA PHAN
EBBA ANDRÉN

Supervisor: Maja Bärring, Department of Industrial and Materials Science
Examiner: Professor Johan Stahre, Department of Industrial and Materials Science

Bachelor's Thesis 2019
Department of Industrial and Material Science
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: A dashboard created with a data analytics tool.

Typeset in LaTeX
Printed by Chalmers Digitaltryck
Gothenburg, Sweden 2019

Abstract

Industries all over the world are adopting the technologies of Industry 4.0 and one of the main reasons for this is to meet the ever-growing demands of the market. This is happening in a Swedish manufacturing company, SKF Group AB, who is the case company of this thesis. Within their World-Class unit, a vast amount of technologies has been implemented and one of these is the information technology. With this information technology, there is a massive influx of data created and stored on a daily basis. With this influx of voluminous datasets, the frequency of the streamed datasets needs to be processed in timely manner as well. In order to fulfill this need of processing and analyzing data, the initial steps moving forward is the “Creation of a Local Data Analytics Center”. The purpose of this center is to utilize the manufacturing capacity, efficiency, flexibility, and as well as increase the global competitiveness by making decisions supported by data. Therefore, as there is a need to process and analyze the data in timely manner, a prompt solution is required. First and foremost, prior to the implementation of a Local Data Analytic Center, this thesis investigates the existing data analytic tools available in the market today and what the functional requirements are of these data analytic tools.

As different companies have different needs, the results of data analytic tools suggested in this thesis needs to be adjusted to suit the requirements and needs to the company of question. However, the methodology applied in this thesis could be adapted to identify existing data analytic tools and the functional requirements of it. There are many different data analytic tools and methods available from different sections and industries, with a wide range of different requirements and levels of knowledge and expertise. Compared to the more traditionally focused data analytics utilized by experts (which requires a high level of knowledge), the delimitations of this “Creation of a Local Data Analytics Center” thesis, has been aligned with different aspects that were important to the case company of this thesis.

The results in this thesis are of importance as they serve in representing the first step taken in order to move forward, closer upon the “Creation of a Local Data Analytics Center”. However, it is important to acknowledge that these results are not flawless and should be critically viewed as there were limitations in the verification step throughout the methods conducted. There are many different aspects to be taken into consideration upon creating a “Local Data Analytics Center”, as it is a vast and complex world of integrated technologies. Therefore, this study only covered a miniscule area revolved around big data analytics and will not investigate the costs and system requirements. This study will be limited to answering the following research questions: “What Data Analytic tools are available on the market?” and “What are the functional requirements of the data analytic tool?”.

Keywords: Industry 4.0, Smart Manufacturing, Manufacturing, Internet of Things, Digitalization, Data, Big Data, Data analytics, Data analytic tools.

Acknowledgements

We would like to start by thanking our supervisor Maja Bärring, who in an engaging manner, read and discussed the content of this thesis and who provided us with very good input. We would like to thank our examiner Johan Stahre for being a part of this journey.

We would also like to thank our supervisor from SKF Group AB, Anneli Sundblom for giving us the great opportunity to work with you and for the support you have provided us. We also want to thank everyone from SKF Group AB, for the openness and support you had towards us.

Last but not least, we would like to thank family and friends who supported us in this journey.

Christina Phan and Ebba Andrén, Gothenburg, May 2019

Contents

List of Figures

List of Tables

List of Abbreviation

BD	Big Data
BDA	Big Data Analytics
CPS	Cyber-Physical Systems
ERP	Enterprise Resource Planning
IoT	Internet of Things
IP	Internet Protocol
LDAC	Local Data Analytics Center
RE	Requirement Engineering
SKF	Svenska Kullagerfabriken (Swedish Ball Bearing Factory in English)

1
Introduction

This chapter starts off by introducing this thesis by describing the background to the context of this study followed by information regarding the case company, SKF Group AB. Consecutively it is followed by further description of the: aim, research questions and delimitations. At the end of the introduction section, the structure of this thesis will be presented.

1.1 Background

Many successive technological developments and innovations within manufacturing processes has been successful in the last few years. These successful technological advancements have brought forth consequences, by putting the global industry on a threshold of the new industrial revolution, Industry 4.0 (Schmidt et al. 2015). One of the world's leading companies specializing in bearing and seal manufacturing, SKF Group AB, aim to utilize their manufacturing capacity, efficiency, flexibility as well as their global competitiveness by entering the next generation of manufacturing technology (Industry 4.0). By doing so, a new production unit has been implemented. This new production unit, which goes by the name of World-Class, is the first of its type to be operated within the company as a unit fully automated and adapted to manufacturing of the spherical roller bearings (SKF 2017).

With the implementation of the World-Class unit, massive sizes of data are collected and stored, and there is a need to make the best of it. With this, there is a demand to make use of the data and that, to be done through the use of a data analytic tool. The analysis of the data can be executed in the analytic aspects of, e.g., organizing, representing, describing, evaluating and interpreting data in order to work more efficiently (Barham 2017). Hence, a study regarding the data analytic tools is necessary and therefore is the purpose of this study. This study consists of investigating what data analytic tools are used and what the functional requirements are for these tools, which is a part of the "Creation of a Local Data Analytics Center". Thus, this study conducted by the thesis workers are in collaboration with SKF, serving as the case company.

1.1.1 SKF Group AB

In 1907 Gothenburg Sweden, Sven Wingquist invented the world's first double-row self-aligning ball bearing and is one of the founders of "Svenska Kullagerfabriken" (SKF), known as SKF Group AB today. SKF Group AB is one of the world's leading companies who specializes in bearing and seal manufacturing (SKF Group AB, 2019). Over the many years of primarily specializing on bearings, SKF has grown and now offer products and services from all five technology (SKF Group AB, 2016) areas such as: bearings and units, seals, mechatronics, services and lubrication systems (SKF Group AB, 2019).

SKF's continuous technology development focuses on the improvement of efficiency and reduction of energy losses, which are some of the many key elements in order to reduce their environmental impact both in their operations and for their customers (SKF Group AB, 2016). In order to make this possible, SKF has now entered the next generation of manufacturing technology (Industry 4.0), through its fully automated and adapted manufacturing of the spherical roller bearings. This manufacturing unit, which goes by the name of World-Class, is the first of its type to be operated within the company (SKF Group AB 2017). Their focus is to ensure a high-quality performance in order to serve their customers on a global scale in the best way possible (SKF Group AB, 2016). In the World-Class unit, a collection of massive data is collected and stored continuously. Thus, there is a demand to make use of the data through data analytic tools where the analysis of the data can be executed in different aspects in order to work more efficiently. For that reason, there is a demand to study the data analytic tools. This study investigates what data analytic tools are used, and what the functional requirements are for these tools as they will be a part of the "Creation of a Local Data Analytics Center". This study is in collaboration with SKF, serving as the case company.

1.2 Aim

With the manufacturing industry becoming more digitalized, in this case SKF, one of the consequences is that massive data sets are collected. In order to make use of the data, there is a demand for the "Creation of a Local Data Analytics Center", whereas a data analytic tool will be part of. The data analytic tool has to be able to perform in the data analytic aspects, e.g., interpret data, to enable the case company to work more efficiently. In order to work more efficient and independently, the idea is to enable the opportunity of conducting some data analytics, preferably locally and quicker by the employees on site within the factory. Therefore, the aim of this thesis is to identify what data analytic tools are available on the market and what the functional requirements are for the data analytic tools.

1.3 Research questions

The following questions are to be answered in this project:

- What "Data Analytic-tools" are available on the market?
- What are the functional requirements of the data analytic tool?

1.4 Delimitations

There are many different aspects to be taken into consideration upon creating a "Local Data Analytics Center" that is necessary to investigate. To fit this thesis context, the study will consist of investigating the available data analytic tools and its functional requirements. Furthermore, the verification of the functions for the data analytic tools are conducted through available information, in forms of demos and tutorials, originated from the company themselves or through video-sharing websites. This is due to it not being feasible within the given timeframe, waiting for access the tools, then test and verify each and every one of them. The costs and system requirements will not be investigated, but only what tools are available and what the functional requirements are for the data analytic tool.

1.5 Structure of the thesis

The structure of the thesis initially starts by presenting the theoretical framework in Section 2. Section 2 covers the theory behind Industry 4.0, IoT, and Big Data (with some sub-categories involved). Following this is Section 3, which focuses on the chosen methods conducted in order to potentially answer the research questions stated previously in this thesis. In Section 4, the results of the studies conducted are presented. Section 5 discusses the results presented and lastly but not least, the conclusions and recommendations are presented in Section 6.

2

Theoretical framework

In this section, the theoretical framework which will be the foundation of this thesis is introduced. It covers the following topics: Industry 4.0, IoT and Data and Information. The Data and Information in section 2.3., covers the following sub-categories: Big Data, Sources of Data in the Production System, Data Analytics and Requirement Engineering.

2.1 Industry 4.0

Advancements within the manufacturing processes have been achieved through successive technological developments and innovations (Schmidt et al. 2015). These technological successes have, on a global scale, put the industries on the threshold of a new industrial revolution: *Industry 4.0.* (Popkova et al. 2019). The first time Industry 4.0. was introduced, was in an article published by the German Government in November 2011. Their concept of Industry 4.0. is a result from the initiative regarding a high-tech strategy for 2020 (Zhou et al. 2015). However, the term Industry 4.0. (the fourth industrial revolution) evolved from all of the previous industrial revolutions, in which describes a system that enables gathering and analyzing data in an automated way. The reason behind this is to realize intelligent decision making (Bunse et al. 2014).

The general idea of Industry 4.0 is to integrate a number of different technologies in order to improve the efficiency and responsiveness of an industrial production system. This is realized by combining some key elements of the following: the emerging new sensor technologies which can be used to gather data, improved technology in terms of the capacity to store the data, and by increasing the possibilities of analyzing massive amounts of data. With this, the accessibility of the data and what it represents in the analyzing process of trends and behaviour, makes it possible in supporting and alleviating the way decisions are made today. Some of the different technologies that correlate to this complex technological system are: Internet of Things (IoT), Big Data (BD), Cyber Physical Systems (CPS), Machine Learning, Additive Manufacturing, and Robotics to name a few (Ahuett-Garza et al. 2018). In the following sections, more information regarding *IoT* and *Data and Information* (related to BD) will be looked into.

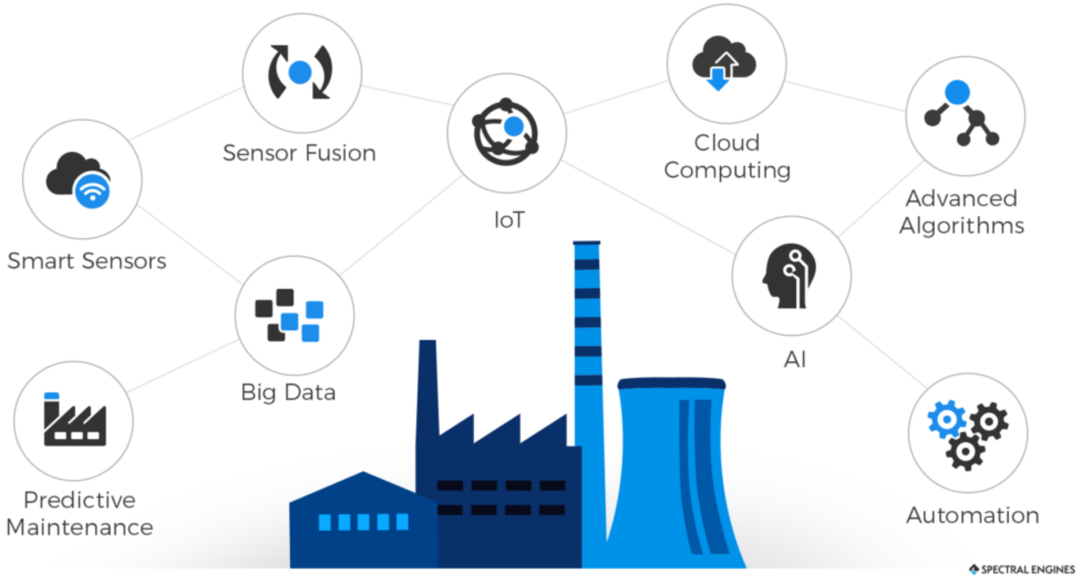


Figure 2.1: Different technologies that correlate to this complex technological system of Industry 4.0. (Spectral Engines Oy, 2019) Reprinted with Permission.

2.2 Internet of Things

In 2010, the earth's human population was surpassed by the number of Internet connected objects (Al-Fuqaha et al. 2015). In a world where billions of objects are connected in some way, to be able to communicate, sense, and share information through interconnection over a public or private Internet Protocol (IP) networks, is the world of *Internet of Things* (IoT) (Patel et al. 2016). Atzori et al. (2010) describes IoT as "a world-wide network of interconnected objects uniquely addressable, based on standard communication protocols". However, IoT is not only made of a network of computers, but also of physical objects. It has evolved into a network of different devices of all type and sizes, which are all connected, and are able to communicate and share information. The devices can be, e.g., smart phones, vehicles, people and industrial systems (Patel et al. 2016). IoT enables these physical objects to hear, think, see and perform jobs by having them "talk" with each other (Al-Fuqaha et al. 2015). Through the communication, it enables them to share information and coordinate decisions in between. By doing this, it allows us to uniquely identify billions of devices and also control devices through the Internet remotely (Gubbi et al. 2013).

There are many challenges with IoT that must be addressed to realize the vision. Some key examples of these challenges are: security, performance, reliability, and availability to name a few (Al-Fuqaha et al. 2015).However, as IoT in the future will link the digital and physical entities together by the means of information and communication technologies (ICT), it will enable a new class of services and applications in many areas (Miorandi et al. 2012). To summarize, the main thing about IoT is the integration of several communications and technology solutions (Atzori et al. 2010). By providing intelligence for planning, decision making and management, these interconnected objects are able to collect data, analyze the data and utilize it to instigate action (Patel et al. 2016). IoT is not an individual system but should rather be seen as an integrated infrastructure upon which many services and applications can run (Stankovic 2014). In a few years, the IoT will enable new applications to bridge diverse technologies by connecting physical objects in support of intelligent decision making (Al-Fuqaha et al. 2015).

2.3 Data and Information

The magnitude, growth and share of data generated is increasing immeasurably in numerous industries and sectors (Agarwal et al. 2014). One of the things that makes data so exceptional, is in its considerable dexterity in being copied and/or combined with other data, meanwhile being accessed and used by many at the same time. Another thing that makes data so unprecedented, is when the data is, e.g., processed, organized, interpreted or presented as to make them consequential, information is created whereas the information provides context for the data (Manyika et al. 2011). The increasingly available data and information was made possible due to the successful technological advancements made. However, with this comes the challenges as to how to, e.g., handle or analyze the data with efficiency (Chen et al. 2014). As data and information is of significance, the following sections will further explore *Big Data*, *Sources of Data in Production Systems* and *Data Analytics* as they are fundamental terms that fit in this context.

2.3.1 Big Data

The term *Big Data* (BD) has no unified definite definition as BD is an abstract concept and people have different opinions on what it means (Barham et al. 2018). However, the importance of BD has been globally recognized in many different fields and sectors. In the article written by Chen et al. (2014), they claim that the term BD is mainly used to describe the prodigious datasets. With the growth of large datasets drastically ascending, some challenging problems surge which demands prompt solutions. The challenges expressed by Chen et al. (2014) with handling BD are the following:

- The *collection and integration* of massive data from different data sources due to the advancements made within Information Technology (IT).
- The *rapid growth of data* challenges the requirements from a hardware and software infrastructure in companies. The capacity of the existing IT architecture and infrastructure face challenges expressed in terms of, e.g., safeguarding, accessing, storing and managing.
- Effectively *access* the datasets at different levels from different process stages (during analyzing, modeling, visualizing and forecasting) to support and improve decision making.

To reach the ultimate goal of taking advantage of data analytics, the core of big data processing encompasses in the stages of collecting, storing and transporting. Big Data, perceived from a data analytics perspective, is defined by the four V's – Volume, Velocity, Veracity, and Variety. In order to classify a problem as a Big Data problem, it is assumed that all or parts of the four V's must have its requirements met (Zhou et al. 2014). What this means is that the four V's in BD still brings many challenges, which requires prompt solutions and should not be undervalued (Buhl et al. 2013). Zhou et al. (2014) describes the four V's and its challenges as following:

- *Volume* refers to the amount of data generated and stored which challenges the current information technologies.
- *Velocity* represents the frequency of streamed data that needs to be processed in timely manner, which cannot be limited to time consuming traditional methods.
- *Veracity* refers to the legitimacy and quality of the data as it can be challenged by the discrepancies in the data collection.
- *Variety* represents all the different types of data, e.g., structured or unstructured data.

With the appearance of internet-based companies, the many traditional process-oriented companies are challenged. In order to stay competitive in the market, the companies now need to become a part of this paradigm shift and become a more knowledge-based driven by data rather than by process (Zhou et al. 2014). As the focus of this thesis revolves around the research questions mentioned prior, the answers may possibly serve as a support to the collaborating company in the selection process of which or what tool they would like to further delve into. The tool may eventually serve as a solution, which potentially may improve the Velocity, Veracity and Variety of data, in order to support and alleviate in better decision making driven by data.

2.3.2 Sources of Data in the Production System - ERP

The extensive utilization of many different advanced management systems has improved the information and management level of enterprises and is continuously doing so (Barham 2017). The sources of data in a production system may come from a variety of sources, e.g., external reference systems and Corporate Business Systems. Data gathered from an external reference system is data that is related to new machinery, in which may be provided by the machine tool manufacturers. However, a majority of the operational data is sourced from *Corporate Business Systems*, e.g., *Enterprise Resource Planning* (ERP), as it hosts and provides key operational manufacturing data such as machining times and more (Skoogh et al. 2012).

At present, the ERP systems are intractable in adapting to the changes happening in the processes within an organization. The ERP system is required to promptly adjust to the changing processes and value-added chains in order to streamline the internal structure in the organization (Babu et al. 2014). Besides the challenge mentioned by Babu et al. (2014), Skoogh et al. (2012) presents an additional numbers of challenges brought forth by the various data sources combined with the difficulties in the data collection process. These challenges are described in the following list:

- *Data accuracy*: Due to the discrepancies in the data collection caused by the variety of internal and external sources, the accuracy of data may need further confirmation. Hence, an investigation is required to investigate the source and the format of collected data to ensure the legitimacy and quality.

- *Data correctness*: It is important to question if the collected data is correct, as communication problems may have occurred during the data collection process. An example is if the automated or manual data collection system could have possibly labeled certain events incorrectly.

- *Data duplication*: The same type of data may have originated from different sources. An example of this occurrence is when the machining time of a specific job can be found from the ERP system, meanwhile it is also possible that the machine operators have their own records of the machining time, or more records of the specific job done in a more accurate, detailed and updated data than the ERP system. Therefore, upon the data duplication occurrence, human involvement is necessary in order to judge which of the two would be the most reliable source of data.

- *Data consistency*: As the same type of data may be sourced by various systems, the data describing the same event in a model may also vary. Hence, if there is a notable difference, the model builder's judgement is necessary in order to judge which would be the most reliable source of data. Upon this data inconsistency, further data analysis is required in order to identify the cause.

- *Data timeliness*: Depending on the modelling purpose, different data from different sources may be required. As the model may live through several iterations in both system design and system management when used for production systems development, the data used may be collected from multiple occasions from various systems.

- *Data validity*: Due to discrepancies, inconsistency or other reasons, the collected data may not fully represent the real-world system. Validity checks of the data are therefore required, as the data often only are "samples of historical event".

- *Data reliability*: The reliability is similar to the validity, however "refers more to being reliable in the eye of the model builder and other stakeholders".

- *Data completeness*: Not all necessary data can be found in the existing sources. Therefore, additional data collection or even assumptions from process experts may be necessary in order to make the whole data collection complete.

In the ERP system, the data is increasing in massive volumes and this exposure to BD where its combined analysis of large amounts of structured and unstructured data from different systems, is executed within a short amount of time. Due to its limitations, a better usage of predictive analytics is required in Big Data Analytics (BDA) in order to uncover concealed patterns and the connections to enable visualization and exploration of data (Babu et al. 2014).

2.3.3 Data Analytics

Big Data Analytics (BDA), refers to the usage of BD to understand and transform the influx of data into usable information and generate insights to aid in better decision-making. Therefore, it can also potentially allow right decisions to be made at the right time based on quicker, more factual, more efficient and more effective gathering and analysis of different sources of data (Barham 2017). The progressions of BD and predictive analytics have paved new paths for exploring new boundaries in analytics-driven automation, as well as in aided decision management in regards of front-line operational decisions (Babu et al. 2014). The key in developing a valuable BDA lies in the vision and determination from the organization. In order to take this step, the organization must have the commitment, imagination and courage as it is a challenge to handle everything that are part of BD (Davenport, 2014).

At first, it may seem as if the primary value from BD comes from the data itself. However, that is not the case. The processing and analysis of the data and the services, insights and products that emerge from the analysis is the prime value regardless of how big the data is in its raw form. In addition to this, there is more to BD than just the software, hardware and the data itself. The people in the organization who are responsible for the utilization of BD, hold an important role as well. The human aspects considered can either be described as vertical or horizontal data scientists. The vertical scientists have great knowledge in the field of question, though very limited. Whilst the horizontal scientist combines the vision with the fundamental knowledge of different technologies (Davenport 2014).

2.4 Requirement Engineering

In order to identify some requirements, Requirement Engineering (RE) can be used (Nuseibeh et al. 2000). The term itself has many definitions. However, the common idea is to define the requirements by identifying what the users want from a computer system and to understand the needs in terms of design (Sutcliffe et al. 2012). With other words, RE covers the process of developing, identifying and verifying the system requirements (Hersman et al. 2010). The activities involved in RE differs as it is dependent on the project and its needs (Sutcliffe et al. 2012). The suggested process map in *"Chapter 18- User-Centered Requirements Definition"* written by Sutcliffe et al. (2012), has been adapted to fit this specific thesis project. The full descriptive adaption of the processes can be found under Section 3.3.

3
Methodology

The methods in this study has been adapted to this thesis context. The main idea has been to collect information and have its contents analyzed, subsequently conclusions were made. Throughout the thesis, literature studies have been conducted from start to finish. The literature studies were conducted through using a number of databases such as: Ebsco Host, IEEE Xplore Digital Library, ScienceDirect, Scopus, SpringerLink, OAlster and ProQuest. Some other sources providing journals such as Science. The procedures and methods conducted for this thesis is presented in this section.

3.1 Procedures

The context and aim of the thesis required an exploratory approach and the reason for that is because of its unusual field of research. The exploratory approach chosen for this thesis consists of a qualitative research methodology as this approach assist in the progress of gathering information. To characterize the link between the theoretical framework of scientific research and its methodology with the context of the project being researched, there are some things that must be investigated and taken into consideration (Bryman & Bell, 2011). The first consideration is to decide which theoretical frame of reference to process whilst the second is to investigate what data collection is required to analyze and build theories upon (Creswell 2013).

During the data collection phases, data is broken down into smaller codes. By doing this it enables the thesis workers to grow their understanding and perspectives, shaping the code on the context of the project. This methodology, open coding, refers to the coding practice of data retrieval done from, e.g., interviews. The interviews are further broken down, examined, compared, conceptualized and categorized in order to be put into a specific context and sorted in categories (Bryman & Bell, 2011)

An overview of the procedures is illustrated in Figure 3.1. In this figure, a literature research was initially conducted. Following this, the whole Benchmarking and Interview process (planning, benchmarking/interviewing, process data and data analysis) was conducted alongside of each other. At the data analysis stage, if the data collected was insufficient, the thesis workers would go back to the benchmarking or interviewing stage for further investigation. When the initial data analysis stage is complete, a final data analysis would be conducted in order to present the end results.

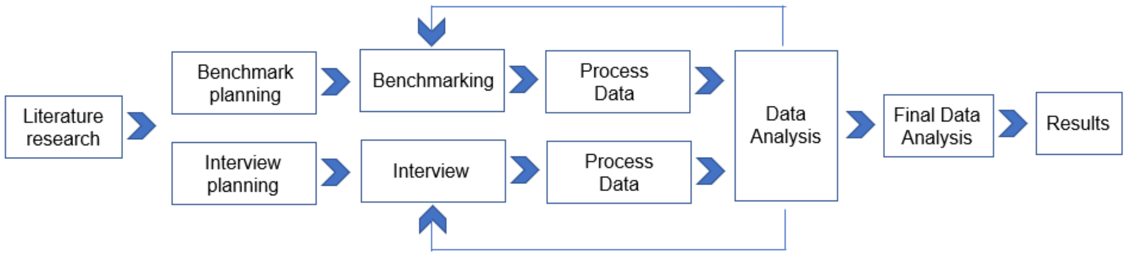


Figure 3.1: Overview of the procedures.

The procedures illustrated in the figure 3.1, includes six different steps conducted. The relationship between the procedures and the six steps are illustrated in Figure 3.2 below. The six steps; Context, Plan, Interview/Benchmarking, Documentation, Analyze, and Results, are further described as the following:

- **Context** - The thesis is initiated by identifying the problems and gathering information in order to get a better understanding and describe the purpose of the study and formulate research questions. This is followed by the next step, planning.
- **Plan** - Planning of how to and what information is required to answer the research questions, both through literature and people involved. Here, the interviews are planned depending on how and what information is of interest based on the research questions and the context of this thesis.
- **Interview/Benchmarking** - When the planning phase is done, it is followed by conducting the interviews which was recorded and transcribed.
- **Documentation** - The interviews were further documented through the transcriptions and notes.
- **Analyze** -The documentation was further analyzed by a methodology of a "Live mind-map" where keywords were identified and categorized. With this, some requirements emerged, however was deemed insufficient and more information was required in order to complement the initial requirements. To further identify some requirements, RE is used. An adapted version of RE has been used in order to fit this project. The end result of this is the creation of Requirements of Specification. The Requirements of Specification was verified with the supervisor of interest in this thesis.
- **Results** - Presentation of the results.

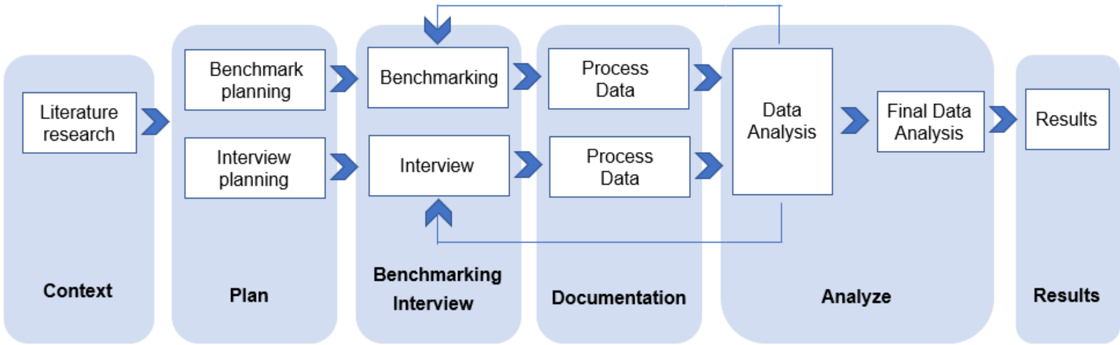


Figure 3.2: Overview of the procedures included with each of the steps.

3.2 Qualitative study

The qualitative study conducted for this thesis has been adapted to fit the context (Creswell 2013) and will be further described in this section. The study is initiated by identifying the problems and gathering information in order to get a better understanding and describe the purpose of the study and formulate research questions (found in Section 1.3). This is followed by planning how to and what information, both through literature and people involved, is required to answer the research questions. Interviews are planned in ways depending on how and what information is of interest based on the research questions and the context of this thesis. When the planning phase is done, it is followed by conducting the interviews, which in turn was recorded and transcribed. The transcriptions were further thematically analyzed by a methodology of a "Live mind-map" where keywords were identified and categorized. The interviews and live mind-map are all a part of the method Requirement Engineering, which is used when identifying requirements. An adapted version of RE described by Sutcliff et al. (2012), has been used in order to fit this project. The combination of these methods is the groundwork of the adapted Requirements Specification which was used in this thesis. A market research (benchmarking) of the existing data analytic tools was conducted in which a list was created. The list of the existing tools and the Requirements Specification were combined into a deconstructed version of the Elimination Matrix proposed by Johannesson et al. (2013), which was executed in three-steps to find out which of the tools are most feasible when implemented into a LDAC.

3.3 Adapted Requirement Engineering

The Requirement Engineering method, described by Sutcliff et al. (2012), has been adapted to fit the context of this thesis. This method involves four steps: *Elicit, Analyze, Model* and *Validate*. Each of the four steps, now adapted, are described in the following consecutive order:

1. *Elicit requirements step*: The initial step of identifying the requirements starts by combining objectives with facts, information and goals expressed by the end-users. This was conducted through interviews, where the interviewees can express inten-

tions and their requests or wish lists.

2. *Analyze requirements step*: In the second step, the preliminary information gathered from previous step, is organized and linked between facts. As an example, when an objective is identified relationship is added depending on what arrangements needs to be acted upon and its implications for the people involved (e.g., actors, organization unit) in order to achieve the goal.
3. *Model requirements step*: The third step is entwined with the analyze requirements as the analysis produces the models, that document the facts and their connections. The goal of the modeling is to create context on the implications and changes. As it is not always easy and clear in understanding the goals without describing the process on how to possibly achieve them and the relationships it has to people and processes, this step is therefore essential.
4. *Validate requirements step*: With the previous steps working as the foundation, a discussion is held with the end-users in terms of validating and negotiating of requirements. The requirements are either agreed on, rejected or resolved among the people involved.

The steps 2-4 should be done in an iterative loop as it is rare for everything to be done correctly the first time. It is recommended that this should be done despite the fact that the first requirements being formally agreed upon by all parties (Sutcliff et al. 2012). By doing so the chances of improving the quality of the requirement specification increases (Stal 2014), as other requirements may surface through the iterative loop of step 2-4 (analysis, modeling, validation, and negotiation) (Sutcliff et al. 2012). Therefore, the requirements may be an adequate foundation when designing and implementing the system (Stal 2014). Following this, the goals then break down to form a desired descriptive state of the system. Some of the implications it may bring could affect the management (e.g., organization, resources, staff activities) whilst some become functional requirements (Sutcliffe et al. 2012).

3.3.1 Interviews

Semi-structured interviews was used in this thesis, which is one of many forms of research interviews. In a semi-structured interview, a set of defined topics and questions are available to the interviewer. Depending on the answers given by the interviewee, some further topics and or questions may be included as it may be relevant to the context of the project. The order of the predefined questions asked does not strictly follow any specific order at all times as it may be adapted to the flow of the interview session itself (Bryman et al. 2011). The interviewee is asked about their experiences in order to understand how they think and feel about something. Hence, the questions in between the interviews are free to vary, as it is likely to change substantially between participants (Fylan 2005). The flexibility of semi-structured interviews makes them well suited to answering the “why”-questions. By adapting the questions and areas of interest depending on the interview session, it makes it possible to further address the prior missing aspects that turned out to be significant to the individual participants. By doing so, it is possible to gain a better understanding of the research questions.

Prior to interviewing, there was some guidelines for research ethics that were taken into account. These ethics were described by (Bryman et al. 2011) and aligned with the formation of the interview. The formation of the interview was formed in a way, in honour of the ethics, by not revealing the interviewees' name or title. This gave the interviewee the choice to either be anonymous or not with comfort and trust. In addition to this, the interviewees were informed of the interview being recorded. However, prior to recording, the interviewee had to give their full consent through an oral agreement before proceeding. The general framework of the questionnaire of this thesis, is divided into four sections: introduction, information, acceptance and summary. This is to alleviate the reviewing and analysis of the qualitative study. The four sections are described in the following list:

1. *Introduction*: An introduction was given to the interviewee and they were introduced to the project. The interviewee was also informed of the honouring aspects of the ethics and informed on how the collected data would be used. In addition to this, an oral agreement on the recording was either confirmed or denied.
2. *Information*: The interviewee is given further information on the context of the project, the purpose and of the importance of their role in the project. The knowledge of the interviewee regarding the context was assessed, in order to give the participant more information if needed.
3. *Acceptance*: The interviewee was assessed regarding their acceptance in regards to the current situation and the scope of the project. In order to increase the acceptance of the implementation of an LDAC, the interviewee was asked of their needs and requirements in order to freely and willingly use it.
4. *Summary*: Here feedback was given by reviewing the session and confirming important points stated during the interview. To end the session the interviewee was once again informed of how the collected data would be used and how they could access the end report.

An initial questionnaire was created, and the first interview was conducted, reviewed and transcribed. Reviewing after the interview made it possible to refine the questionnaire prior to the next interview. This was done in order to increase the quality of the questionnaire, which in turn may increase the quality of information gathered which is beneficial to the context of this thesis. Figure 3.3 illustrates this interviewing process. In order to be more present in the interview and focus on the interviewee, the interviews were all recorded. The recorded interviews were then reviewed and transcribed in order to further document and eventually identify the needs, challenges and requirements. When all of the interviews had been conducted, all of the recordings fully transcribed and documented, the data collected from the interviews were analyzed which further on provided insightful results. The full questionnaire can be found in Appendix A.

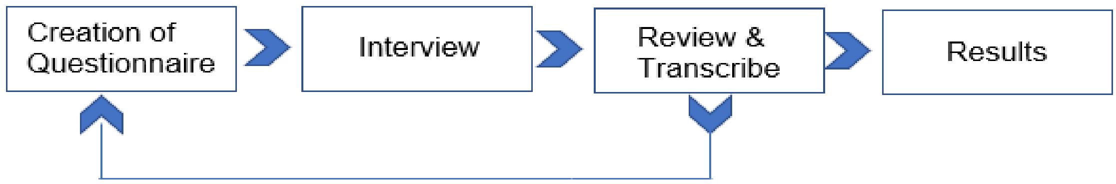


Figure 3.3: Illustration of the whole duration of the interviewing process.

3.3.1.1 Current State Analysis

The factory's current situation regarding data analytic tools was created based on the results of the interviews in order to create a current state analysis. Through the conducted interviews with representatives from many different departments, a holistic view of the needs and level of knowledge when suggesting a list of data analytic tools for the LDAC in this project was depicted. The interviewees were those who would be a part of this project at an early stage, e.g., as an end-user, and also those who potentially would have access to the expansion of the project at a later stage. The interviewees were initially recommended by the supervisor from SKF, as those representatives can give valuable feedback and input due to their work experience or even recommend another person to interview. The interviews were evaluated, then further investigated in order to define the possibilities to achieve accessible tools for the LDAC, whereas the goal is to improve the use of data analytic tools and hence data analysis.

3.3.2 Live Mind-map

In order to analyze the transcribed interviews, the thematic analytic method was used which focuses on identifying the themes and patterns of the answers on the questions asked during the interviews. The steps taken upon using thematic analysis, is to firstly collect data (Aronson 1995). In this case, the first step is transcribing the interviews and then gather the data, which is done by reviewing the transcribed documents. Following this is the second step, where data is identified (keywords) and already related to a specific pattern (Aronson 1995).

These steps described in the thematic analysis process, is conducted through a live mind-map, with some additional steps. The reason behind a live mind-map is to hold discussions whilst identifying and categorizing the keywords. This thematic analytic method is the foundation of the resulting live mind-map conducted. The data collected from the interviews went through the data analysis process, where keywords were identified and written on notes (post-its). The post-it notes, were initially placed depending on its context. However, as more post-it notes were placed, they created their own respective group hence the themes were identified. Whilst organizing and thematizing, categories were further identified. During the whole process of the live mind-map, dialogues were held between the thesis workers regarding the position of each and every one of the post-its, in which could change throughout the process, due to their positions being questioned and built with valid arguments.

3.3.3 Requirement Specification

By designing a Requirement Specification, the requirements for a data analytic tool were identified through the methods described in previous sections (with its foundation built based on RE). The purpose of the requirement specification is to define a number of criteria in which functional requirements and non-functional requirements are listed. The functional requirements describe the functions that the data analytic tool must have to fulfill the users needs and the non-functional requirements describe the product's safety- and usability requirements (Projektledning 2019). When the criteria are made, the Requested (Ö) criterion priority score will be scored using an interface matrix, which is further described in the following section.

3.3.3.1 Priority score of the requested criterion

The requested criterion (Ö) score was based on using an interface matrix. On the requirement specification, a priority score of 1-5 was distributed where a score of 5 is the most important request criteria to fulfill. The Request criterion was distributed a score depending on how important they were in order to reach the goal of this thesis context, which is to identify the functional requirements for a data analytic tool. The request criterion is founded by expressed requests or wishes by the participants, of whom participated in the interviews. These requests are meant to be used on the purpose of potentially identifying further functional requirements for a data analytic tool, which would challenge the many different data analytic tools in terms of functionality following the benchmarking step.

Each of the requested criterion was compared to each other in an interface matrix, where they were distributed a score of -1=less important, 0=equally important or 1=more important. After all the scores were distributed, a total score was summarized for each of the criteria. The reason as to how or why the criteria were distributed the scores of -1, 0 or 1, are generally explained and is presented in the results section. When the total score is accounted for, a Priority score is distributed among each of the criteria depending on their Total Score. The relationship between the Total score and Priority score is further illustrated in Figure 3.4 below.

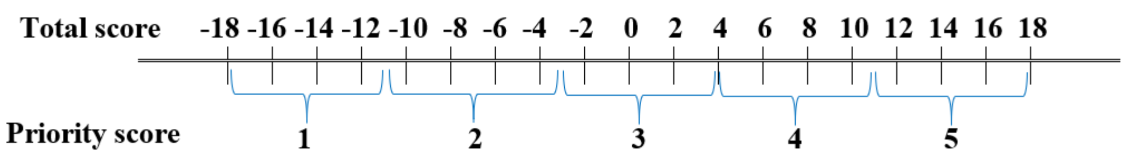


Figure 3.4: The relationship between *Total score* and *Priority score*.

3.4 Market research of analytic tools - Benchmarking

A market research of existing data analytic tools was conducted. The benchmarking on the tools was based upon some of the requirements stated by SKF. Some of the requirements stated is that the data analytic tool must have a strong customer base in order to have something to stand upon. As an example, that the data analytic tool is used by other known companies. The market research emanated from this and based on this, the following three points were taken into consideration:

1. *The customer base*: The purpose of this was to take into account of if the data analytic tool had any major companies as customers.
2. *The reputation*: While researching and finding information of each tool, to also objectively consider the data analytic tools reputation in regards to usefulness, accessibility, security and more (if such information was available).
3. *The upcoming trends*: The upcoming tools were taken into consideration in order to stay open to other options of which tools could possibly eventually fit in as a solution of this thesis context.

3.5 Elimination matrix and process

To eliminate the analytic software tool that are unable to fulfill the requirements and requests from the requirement specification, an elimination matrix is initially applied. The “traditional” elimination matrix was deconstructed to fit into this context. In the deconstructed version, which have different functional criteria (from requirement specification) listed, are requirements in which the tools must sustain in order to make it through the elimination. This elimination matrix version is a combination of listing each tool in the left column, and of the criteria from the requirement specification listed on the top row. Hence, the layout and idea of the “traditional” elimination matrix is applied but deconstructed to fit this thesis context.

Upon the creation of the elimination matrix, each of the data analytic tools were cross-referenced with the Requirement Specification criteria. This process was executed separately between the thesis workers and the reason behind this is due to taking the differences between each individual into consideration, as how every individual work and think vary which may lead to different results. During the cross referencing, one program at a time was investigated more in depth and researched on. The thesis workers would mark each criterion with X or Y if the criteria are fulfilled, left blank if more information is required and - when unfulfilled. The reason why the thesis workers each had to mark with X or Y is to distinguish who did what. Following this, when all of the tools had been investigated independently, the thesis workers will further discuss and explain where and how they found the information. The deconstructed elimination matrix is the foundation of the next phase, which consists of a three-step elimination process conducted to gradually filter out tools.

3.5.1 Elimination process

In this project, the elimination process consists of three different steps in order to find, score, and rank the data analytic tools. The main goal of the first step in this elimination process is to eliminate the programs that did not fulfill the requirements of 1) handling many different data sources and types (criteria 1.1.) and 2) uphold a certain technical standard of the 20th century (criteria 1.2). By technical standard of the 20th century, e.g., means no Windows XP user layout, slow in importing data and more. In the second elimination step the programs are challenged by having to fulfill the requirement of being user-friendly (criteria 2.9). Furthermore, if the tools fully fulfilled the required criteria or if they also had some criteria that was left blank (means that more information is required), they would not be eliminated. In the third and final step of the elimination, is to score and rank the programs survived thus far. The ranking of the tools consists of summarizing all the X's and Y's, and divided with the total sum, where the total sum is the amount of both Required (K) criteria and Requests (Ö) criteria, multiplied by 2 (for each of the thesis workers X's and Y's). To make it simple, the elimination process is conducted in the following three steps:

1. Eliminate the programs that do not fulfill criteria 1.1 and 1.2 (Handle different data sources and types, and Uphold a technical standard of the 20th century. The reasoning behind this is due to the fact that the data analytic tools must be accessible the end-users of the case company).
2. Initially eliminate the software that do not fulfill criteria 2.9 (User-Friendly) and secondary, eliminate the software that do not fulfill the Required criteria (K).
3. Rank the software survived thus far.

4
Results

This section will present the results of the methodologies described in the previous section. The findings will represent in supporting the answers of the specific research questions. The results of this section are presented in the following consecutive order: Adapted requirement engineering, interviews, current status analysis, live mind-map, requirement specifications, priority score of the requested criteria, benchmarking, and last but not least the elimination process.

4.1 Results of the adapted requirement engineering

The results of the adapted requirement engineering consist of four steps. These four steps are the elicit, analyze, model and validate steps. The results from the elicit, analyze, and model requirement step, which consists of conducting the interviews, analyzing them, and part modeling, the results of these steps are presented in section 4.1.1. With the foundation set, following this is the continued modeling and validate requirements steps. The results of these two steps are intertwined in all of the following result section. The reason for this is because the modeling and validating were frequently updated throughout the whole process.

4.1.1 Result of the Interviews

The interviews were conducted as many times as possible within the given time frame. The chosen interviewees were recommendations from the company supervisor of this thesis and as well as recommendations from the interviewees themselves. The participants, representing the company as employees with different roles, were as well from different departments of the company. The departments in question are the following: Process Strategy, Tivo Rollers, Metrology, Maintenance, Industrial Engineering and Factory Logistics. The reason for interviewing so many different people from different departments is to capture as many different perspectives, views, experience and knowledge as possible. This was essential in providing the context of this thesis with relevant information of interest, which is the foundation regarding the needs and requirements closely related to the collaborating company. The number of participants and the departments involved are represented in Table 4.1 below.

Table 4.1: Departments and number of interviewees participated.

Department	No. of interviewees
Process Strategy	4
Tivo Rollers	2
Metrology	1
Maintenance	1
Industrial Engineering	1
Factory Logistics	1

Each of the interviews lasted around 45 minutes whilst each of the transcriptions took around 80 minutes. When all the interviews had been conducted, all of the recordings fully transcribed and documented, the data collected from the interviews were analyzed which further on provided insightful results. These results are presented in the following sections below.

4.1.1.1 Current status analysis

The interviewed personnel pointed out current sources of errors and problems when using data analytic tools to, e.g., report results and more. The knowledge regarding data analytic tools were very limited and basic, however a majority of the interviewees are keen to improvements, expressing the need to find better solutions and tools regarding this area in order to work more proactive and with efficiency. All of the gathered information was carefully assessed and used later in the project in order to possibly solve the problems and identify improvements.

4.1.1.2 Results of the "Live Mind-map"

A thematic analytic method is the foundation of the resulting live mind-map conducted. The data collected from the interviews went through the data analysis process, where keywords were identified and written on notes (post-its). The post-it notes, were initially placed depending on its context. However, as more post-it notes were placed, they created their own respective group hence the themes were identified. Whilst organizing and thematizing, categories were further identified. During the whole process of the live mind-map, dialogues were held between the thesis workers where the position of each and every one of the post-its, which could change throughout the process, due to their positions being questioned and built with valid arguments.

Finally, 233 pieces of keywords were identified, however do note that multiple keywords of the same type were not taken into account when counting the resulting number of keywords. The themes identified is *Vision (orange), Challenges (purple), Technologies (blue), LDAC (green)* and *Required action (red)*. In these themes, the categories identified is: *SKF, LDAC, Data-Hub* (with two sub-categories of *System* and *Data*), *Functions, Requests, Possibilities* and *Results*. An overview of the identified themes and its categories are illustrated in Figure 4.1. Following this is an overview of the themes, categories, sub-categories and a few keywords presented in Figure 4.2 to give the reader a holistic viewpoint. As an amount of 233 keywords were identified, for clarity, each of the themes will be presented separately in this section as it will be hard to combine and present the full results in one. However, a full combined list of all the identified keywords can be found in Appendix C as well. The results of this section are the foundation for the next section 4.1.2., the resulting requirement specification.

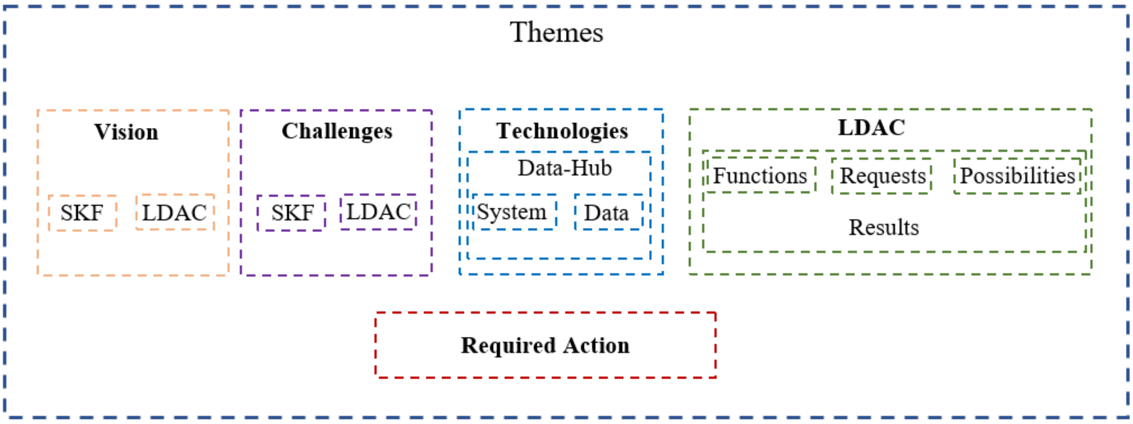


Figure 4.1: Overview of the identified themes and its categories.

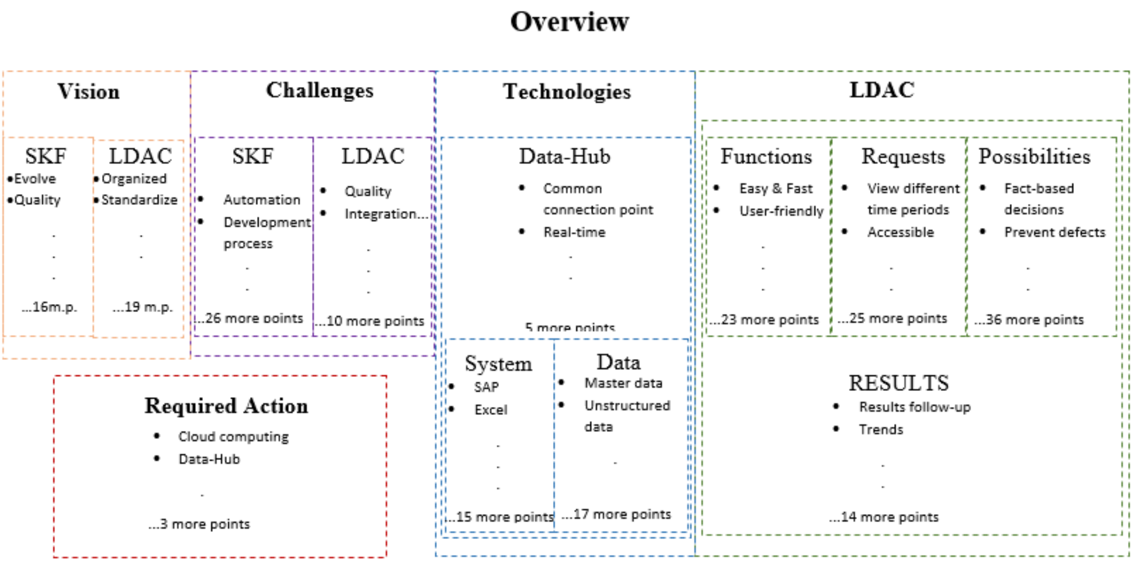


Figure 4.2: Overview over all the themes, categories, sub-categories and keywords.

The theme *Vision* represents the futuristic view expressed by the interviewees, where the views were related to SKF and LDAC. This theme generally depicts the visions of the future. The views could be about, e.g., what they hope to happen or how they want it to look like. The boxes marked in orange belongs to the theme *Vision*, which has the categories *SKF* and *LDAC*. This box has a total of 39 keywords identified, whereas 18 can be found in the category *SKF* whilst 21 is under the category *LDAC*. The results of the orange themed *Vision* is presented in Table 4.2 below. The reason of the theme Vision is due to the identified visions and goals expressed, which implicates the collaborating company and the *LDAC* in different ways, e.g., more standardized and organized work methods or prompt solutions to the challenges Industry 4.0 brings.

Table 4.2: Results of the theme Vision with all of the keywords.

Results of Live mind-map: Keywords				
Theme	Vision			
Category	SKF		LDAC	
	No.	Keywords	No.	Keywords
	1	Want to evolve & improve	1	Organized
	2	Less vertical, more horizontal thinking	2	Uniform picture during meetings
	3	Better data quality	3	Standardize working ways
	4	Bring together several potencies (pull together downpipes)	4	Standardize 2-3 analytic tools
			5	Used by many (receiver)
	5	Reduce measurement time, more efficient	6	Few creators
	6	Less administration	7	Work More Horizontally
	7	Reduce costs in general	8	Keep a clear direction forward
	8	Increase quality	9	Alleviate in daily decisions
	9	High level: IT solutions	10	Make the work easier (manufacturing)
	10	AI	11	Make things easier (people)
	11	Simplify working methods	12	Basis for improvement/changes
	12	Reduce external costs (services)	13	Improvement opportunities
	13	SKF: link machine & cloud (ongoing)	14	Project management/project drive
	14	Transparency for the company	15	Drive projects forward
	15	SKF Center (ongoing)	16	Present good facts
	16	Industry 4.0 for whole SKF	17	Problem solving
	17	Gothenburg & SKF will remain in 10 years	18	Used for a lot
	18	Better at analysis level	19	Prevent/anticipate problems
			20	More time for long-term solutions (no emergency)
			21	Streamline work

The theme *Challenges* represents the things that require more work or that could be challenging to achieve. The challenges expressed are related to SKF and the LDAC. Some of the challenges expressed are: the requirements for a working LDAC might be more than expected, change of working habits and methods, the implications on the units and organizations and more. The boxes marked in purple marks the theme *Challenges*, whereas the challenges has the categories *SKF* and *LDAC* as well. The total of keywords identified is 40, where each of the categories consists of 28 keywords respectively 12 keywords consecutively. The results of this purple category, *Challenges*, is presented in Table 4.3. This theme, *Challenges*, fits the keywords identified of this context. As with changes comes consequences and challenges which in some form or way puts implications on both the company and the LDAC. Some of the challenges could be: change of habits, change how decisions were previously made, challenges of implementing integrated information technologies and more.

Table 4.3: Results of the theme Challenges with all of the keywords.

Results of Live mind-map: Keywords						
Theme	Challenges					
Category	SKF				LDAC	
	No.	Keywords	No.	Keywords	No.	Keywords
	1	Integrate & automate	15	Ability to change	1	SAP & LDAC have access to the same data
	2	Not everything is automated	16	Not so developed	2	Risk: Different data depending on the module (also in SAP)
	3	Administration (produce slides regularly)	17	From manual to digital, complicated interface		
	4	A lot of administration			3	Input data is important (sensitive!)
	5	Requires a lot of work from the beginning, becomes easier with time	18	Work like an engineer with facts	4	Risk with the quality depending on the input data
			19	[redacted]		
	6	Development process of LDAC --> Trial & Error	20	Better connection to the network among all machines	5	Easy way to get data down
	7	Many meetings	21	[redacted]	6	Big data, collects a lot of data from different places --> find connections and abnormalities
	8	Very vertical	22	Support for SAP system		
	9	Focus on one area	23	Data collection should continue despite network problems	7	Combined data from different sources to the LDAC
	10	Different competencies give variations/quality of what is visualized	24	Maybe need a station between station	8	Quality safe
	11	The right conditions for digitalization	25	Digital world? There must be a capacity	9	Data integrated between different departments
	12	Different priorities	26	Has good data, but less good at analyzing		
	13	Challenging to get everyone on the same course	27	[redacted]	10	Integration of data flow
					11	Organize to integrate data
	14	Challenge: consensus in larger group	28	[redacted]	12	Different levels for different skills

The theme *Technologies* covers the topic of what is available and less about what needs to be fixed or what needs to be acquired. Hence, the categories of system and data represents the systems that exist, and the data used. The box located beside the purple one, is the blue one which marks the theme *Technologies*. This theme consists of the category Data-Hub, which in turn has two sub-categories, *System* and *Data*. In this theme, there are a total of 43 keywords. Whereas 7 of them are found under Data-Hub, 17 under System and 19under Data. The results of this blue category, Technologies, is presented in Table 4.4. In this context, technologies is involved in some way or form. In order to effectively use the data analytic tool, a data-hub is required. The reason behind this is as there are many different existing systems with datasets in different data sources, it consumes time and energy to hunt down all of the datasets sought. Hence, the identified general idea is to have a data-hub that creates a common connection point, updated in real-time, to alleviate in the whole process of, e.g., finding data. Worth noting is that the technologies presented in this table is widely used in general within the manufacturing industry. Hence, the specific technologies used by SKF is by no means presented in this table.

Table 4.4: Results of the theme Technologies with all of the keywords.

Results of Live mind-map: Keywords				
Theme	**Technologies**			
Category	***Data-Hub***			
	No.	**Keywords**		
	1	Common data point where everything goes through/is connected		
	2	That the needed data is available		
	3	Build process & work methods around data collection		
	4	More real-time and forward direction (managing KPIs)		
	5	Real-time (maintenance perspective), prevention before something happens		
	6	Data-hub (connect all the different data sources), automated		
	7	Compile the searched data (--> be able to do otherwise, ex. prognosis)		
Sub-category		***System***		***Data***
	No.	**Keywords**	**No.**	**Keywords**
	1	SAP	1	[redacted]
	2	[redacted]	2	SAP
	3	[redacted]	3	[redacted]
	4	[redacted]	4	[redacted]
	5	[redacted]	5	NoSQL
	6	[redacted]	6	SQL
	7	ERP-system	7	Master data
	8	[redacted]	8	Structured data
	9	Visualization of time-based info	9	Unstructured data
	10	[redacted]	10	Data is available in local disk, saved in accessed database
	11	Excel	11	[redacted]
	12	SharePoint	12	Situation & deviations with production volumes (connection)
	13	Ishiawa diagram	13	[redacted]
	14	Kaizen	14	[redacted]
	15	Six sigma	15	[redacted]
	16	[redacted]	16	[redacted]
	17	[redacted]	17	[redacted]
			18	[redacted]
			19	[redacted]
				[redacted]

The theme *LDAC*, cover the expressed functions, requests, possibilities, and results expected or required. The green box, marking the theme *LDAC*, has the three categories: *Functions, Requests*, and *Possibilities*, in which all of these are connected to the sub-category of *Results*. The total amount of keywords identified are 106, whereas 25 of these belong to Functions, 27 belongs to Requests and 38 belongs to Possibilities. The rest of the keywords, 16 of them, is found under the sub-category Results. The results of this green category, LDAC, is presented in the following Table 4.5. The theme LDAC, is the main purpose of the research questions of this thesis. Upon the creation of the questionnaire, this theme was mainly focused upon. Hence, there were keywords identified that fit into each of these categories, representing the functions (e.g., Easy & Fast, User-Friendly), requests (e.g., view different time periods, accessible), the possibilities (e.g., fact-based decisions, early on identify discrepancies), and lastly what the resulting expectations are (e.g., present trends, be able to follow-up on the results).

Table 4.5: Results of the theme LDAC with all of the keywords.

Results of Live mind-map: Keywords						
Theme	**LDAC**					
Category	**Functions**		**Requests**		**Possibilities**	
	No	**Keywords**	**No**	**Keywords**	**No**	**Keywords**
	1	Easy & fast to use	1	When running different types of bearings --> the number	1	Fact-based decisions
	2	Do-it yourself		of transports increases [problems increases	2	Prevent defective products
	3	Easy to learn	2	Be able to shift from day-day & week-months	3	Build away recurring problems
	4	User-friendly	3	Day to day: production volume, production rate > see	4	Work planned and preventive (rather
	5	Create visualization	4	How often the machine stops and the cause	5	Understand connections
	6	Create insights	5	Simple & accessible	6	Everything fits together
	7	Informative visualization	6	Use in factory logistics	7	Better understanding of the result
	8	Create business case	7	Show status in World-Class (on many stations ex AGV,	8	Discuss & create a basis for business cas
	9	Be able to get feedback from X		buffers...)	9	Develop together (common place)
		numbers of years	8	Machine learning	10	Show & discuss
	10	See discrepancies	9	Combine data from different data sources	11	Common place
	11	Deviations analysis in production	10	Static data combined with real-time data	12	Remove screws that were previously stuc
	12	Status update (reconciliation)	11	Database search (for data)	13	Enable cross-references
	13	Charts with different aspects	12	Be able to enter data automatically	14	Challenge old ways of
		(machine- & process data, etc.)	13	Easy connection	15	Support for decisions (Support)
	14	Graphs & Charts	14	Analysis team	16	View, see the same thing --> solve togeth
	15	fast & easy to search for	15	Right people with the right competence	17	Get a connection to implemented
		information/status	16	LDAC -dependent internal	18	Buffer control/status
	16	Show different types of connectio	17	KPI (performance measurement)	19	Spend time on analysis & implementation
	17	Simplification	18	LDAC can reduce administrative workload?	20	See if problems occur at a shift team (ex.
	18	Search-function for example: the	19	More organized		incompetence, wrong working method &
		data you are looking for	20	Clarity	21	Discipline & structured
	19	problem solving tool	21	More independent, less dependent on someone else (to	22	Improvement activities
	20	Click-view		get the data)	23	Collected image presented
	21	Agile (Dynamic)	22	Instructions (LDAC)	24	Important to be to show people
	22	Dynamic	23	LDAC: production-related data (machine & process)	25	Maintain new thoughts and ideas
	23	Collected tools	24	Get the right data (LDAC)	26	Flexible to adapt to recipients
	24	Scale away the unimportant	25	Overview data (machine- , flows-, plant data)	27	Customize data combination for the right
		(simple functions)	26	Simple reporting features	28	Make analysis models central/local (less
	25	History of processed problems	27	Early identify deviations before accidents/serious events	29	Adapt content to the stakeholder
Sub-categ	**Results**				30	Reach out with potential, to the users
	No	**Keywords**	**No**	**Keywords**	31	Create root cause analysis
	1	Result tracking	9	Trends	32	Identify risks
	2	Outgoing quality follow-up	10	Common picture	33	Identify parameters that affect
	3	Reflection on results/Metric (KPI	11	Common entrance > Analyze together	34	Identify problems
	4	The status of the machines	12	Various POV when looking at the data	35	For a selected area, be able to look more
	5	Number of accidents (Without	13	Self-taught/simplification	36	Visually get "Here we burst"
	6	Visualize goals and status report	14	1) Collected Data 2)Analysis 3)Visualize	37	The imagination sets limits in digitization
	7	Visualize performance	15	OEE in real time (now-now)	38	Complete tasks so that it does not cool
	8	Prognosis for the future	16	OEE (Overall equipment effectiveness)		

Last but not least, the red box marks the theme of *Required Action*. In this box, some of the few requirements are identified in order to suggest or recommend actions that is essential to implementing an efficient LDAC. In this theme, 5 keywords have been identified and is presented in the following Table 4.6. This theme was created as the identified keywords, did not fit into the theme of Technologies despite being similar. The identified keywords of the red box is required actions expressed by the interviewees that needs to be acted upon. These keywords were related to either the problems or lack of something that they have experienced, e.g., the IT architecture and infrastructure. As these keywords were considered important, but did not fit into any of the previous themes, it was assigned its own theme, *Required Action*.

Table 4.6: Results of the theme Required Action with all of the keywords.

Results of Live mind-map: Keywords		
Theme	Required Action	
	No.	Keywords
	1	Data-Hub
	2	SKF IT
	3	IT-landscape lacks in general on networks & computers
	4	Upload data > User analysis via web browser
	5	AWS - Amazon Web Services, Cloud

4.1.2 Results of the Requirement Specification

The requirement specification of the data analytic tool has been divided into three categories in order to clarify the requirements defined. The three categories are 1) Tools and options, 2) Layout, and 3) Desired result. The criteria numbers represent the category followed by the number of the same category, e.g., criteria (1.1.) means (1=category.1=number of the category.). The first and second category list the functional requirements whilst the third category list the non-functional requirements.

The Requirement Specification resulted in 40 different criteria. Out of all the criteria, 21 of them are Required (K) whereas 19 of them are Requests (Ö). An overview of the 21 Required criteria are represented in Table 4.7. In this table, these required criteria are from the first (Tools and Options) and second (Layout) category and none from the third (Desired result) category. From the first category, 12 out of 16 requirements are listed. Also, from the second category, 9 out of 11 requirements are marked as Required criteria (K). Table 4.8 represent the overview of the 19 Requested criteria (Ö). In this table, the requested criteria consist of all three categories with the first, second and third criteria with the following number of requirements: 4 out of 16, 2 out of 11, 13 out of 13 of each category respectively.

The full requirement specification, with the Required (K) and Requested (Ö) criteria, is in Appendix D. Upon looking at the full requirement specification, do note that the verification column, which is usually a part of the requirement specification, is left blank. This is due to the limitations of not having full access to all of the tools, hence the thesis workers were unable to personally verify through the tools themselves. In order to verify, a different approach was necessary, and this was conducted through different verification methods. More about the verification methods can be found in the elimination process section 4.3., and the reason for this is due to the various verification methods being fundamental in the elimination process' different steps.

Table 4.7: Overview of the Required criterias (K) in the Requirement Specifications.

Criteria	Description	Criteria	Description
1.1	Handle many different data	1.12	Recorded History
1.2	Technical standard of 20th Century	2.1	Simple
1.3	File options	2.2	Clean
1.4	Edit options	2.3	Easy
1.5	View options	2.4	Clarity
1.6	Insert options	2.5	Do-it-yourself
1.7	Format options	2.6	Fast/Quick
1.8	Data options	2.7	Agile
1.9	Solution tools	2.8	Accessible
1.10	Tools	2.9	User-friendly
1.11	Help Search Tool		

Table 4.8: Overview of the Request criteria (Ö) in the Requirement Specifications.

Criteria	Description	Criteria	Description
1.13	Data search tool	3.5	Enable back-tracking...
1.14	Add-ons	3.6	Data input done automatically
1.15	Filter options	3.7	Combine data from different data sources
1.16	Access to general data-hub	3.8	Static data combined with real-time data
2.10	Dynamic/Flexible	3.9	Use/identify accurate data
2.11	Overview	3.10	Instructions
3.1	How often a machine ...	3.11	Minimize administrative work
3.2	Survey/Overview data	3.12	Simple report / share functions
3.3	Early-on identify...	3.13	Machine learning/AI
3.4	Different levels of analysis		

4.1.2.1 Results of the priority score of the requested criteria

The priority score of the requested criteria was scored using an interface matrix. The resulting interface matrix made of all the requested criteria are presented in Table 4.9. In this table, there are 6 criteria with a priority score 5, 1 with the priority score 4, 4 with the priority score 3, 6 with the score 2, and 2 with the score 1. The criteria with its respective priority score of 1 to 5 is presented in Table 4.10, as each of these criteria have their descriptions presented as well. For clarity reasons, each of these criteria with a short comment given by the thesis workers, are presented in Table 4.11 below.

Table 4.9: Interface matrix of the priority score of the request criteria.

Interface matrix: Priority score of the request criterias																			
	1.13	1.14	1.15	1.16	2.10	2.11	3.1.	3.2.	3.3.	3.4.	3.5.	3.6.	3.7.	3.8.	3.9.	3.10	3.11	3.12	3.13
1.13.		1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1
1.14.	-1		0	0	0	0	0	0	0	1	0	1	1	1	1	-1	0	0	1
1.15.	-1	0		0	0	0	-1	0	-1	1	0	1	1	1	1	-1	0	0	1
1.16.	-1	0	0		0	1	1	1	0	1	1	1	1	1	1	0	0	0	1
2.10.	-1	0	0	0		0	0	0	0	1	0	1	1	1	1	-1	0	0	1
2.11.	-1	0	0	-1	0		0	0	0	1	0	1	1	1	1	-1	0	-1	1
3.1.	0	0	1	-1	0	0		0	0	1	0	1	1	1	1	-1	-1	-1	1
3.2.	-1	0	0	-1	0	0	0		-1	1	0	1	1	1	1	-1	-1	-1	1
3.3.	0	0	1	0	0	0	0	1		1	0	1	1	1	1	-1	-1	-1	1
3.4.	-1	-1	-1	-1	-1	-1	-1	-1	-1		0	0	0	0	0	-1	-1	-1	0
3.5.	-1	0	0	-1	0	0	0	0	0	0		0	0	0	0	-1	-1	-1	1
3.6.	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0		0	0	0	-1	-1	-1	0
3.7.	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0	0		0	0	-1	-1	-1	0
3.8.	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0	0	0		0	-1	-1	-1	0
3.9.	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0	0	0	0		-1	-1	-1	0
3.10.	-1	1	1	0	1	1	1	1	1	1	1	1	1	1	1		0	0	1
3.11.	-1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0		0	1
3.12.	-1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0		1
3.13.	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	-1	0	0	0	0	-1	-1	-1	
Total score	-16	-4	-2	-9	-4	-2	-3	0	-5	12	4	12	12	12	12	-13	-9	-10	13
Priority score	**1**	**2**	**3**	**2**	**2**	**3**	**3**	**3**	**2**	**5**	**4**	**5**	**5**	**5**	**5**	**1**	**2**	**2**	**5**

Table 4.10: Overview over the priority scoring and the requested criteria.

Priority Score	Requested criteria (Ö)	
5	3.4.	Different levels of analysis
	3.6.	Data input is done automatically
	3.7.	Combine data from different data sources
	3.8.	Static data combined with real-time-data
	3.9.	Use/identify accurate data
	3.13.	Machine Learning / AI
4	3.5.	Enable back-tracking of X years
3	1.15.	Filter option
	2.11.	Overview
	3.1.	How often a machine stops and its cause
	3.2.	Survey/Overview data
2	1.14.	Add-Ons
	1.16.	Access to general Data-Hub
	2.10.	Dynamic/Flexible
	3.3.	Early-on identify discrepancies prior to failure/serious event
	3.11.	Minimize administrative workloads
	3.12.	Simple report/share functions
1	1.13.	Data Search Tool
	3.10.	Instructions (LDAC)

Table 4.11: Comments on the priority scoring of the request criteria.

Comments on the score of the request criterias			
1.13.	An option of being able to search for specific data is beneficial, however not more important than some of the other criterion	3.5.	Enable backtracking of X amount of years is beneficial, hence it is either equally or more important.
1.14.	Having Add-ons is beneficial, however compared to some criterias it is equally important or less important	3.6.	For data input to be done automatically is very crucial, as it is not "feasible" to manually enter every data due to many aspects. Therefore, it is either equally or more important.
1.15.	A filter option is beneficial, however compared to some criterias it is equally important or less important	3.7.	To be able to combine data from different data sources is one of many reasons for a data analytic tool. Therefore, it is either equally or more important.
1.16.	Access to a general data-hub is very beneficial, however compared to some it is still less important as it is possible to work around this. It is more important to get some other criterias fulfilled.	3.8.	Static data combined with real-time data is one of many reasons for a data analytic tool. Therefore, it is either equally or more important.
2.10.	If the tool is dynamic/flexible is beneficial, however it is equally or less important compared to some criterias.	3.9.	To use/identify accurate is the very core for data analytics to be used and present accurately. The better quality of data the better quality of data analytics which minimize the risk of false reports. Hence, it is either equally or more important.
2.11.	An overview function is important; however, it is equally or less important compared to some criterias.	3.10.	To have instructions is good to have, however compared to other criterias it is generally less important. Hence the user might require a higher level of competence to be able to do it properly.
3.1.	How often a machine stops, and its cause is good to know, however it is a mix of more, equally or less important compared to some criterias.	3.11.	To be able to minimize administrative workloads is very beneficial in many aspects, however compared to some criterias it is generally equally or less important.
3.2.	A survey/overview over the data is good to have, however compared to some criterias it is a mix of more, equally or less important.	3.12.	To be able to simply report or share functions is beneficial in many aspects, however compared to some criterias it is generally equally or less important.
3.3.	Early-on identify discrepancies prior to problems is good to have, however compared to some criterias it is a mix of more, equally or less important.	3.13.	The use of machine learning/AI was highly requested and could alleviate work in many ways. However not many tools offer this solution - yet. It is however of high importance with a mix of being equally or more important. With the advancements made in data analytic tools, it may be possible in the near future.
3.4.	To be able to have different levels of analysis is very important to this context, hence it is either equally or more important.		

4.2 Results of the Data Analytic tools - Benchmarking

Upon conducting the market research of existing data analytic tools, these tools were not perceived as commonly used within the manufacturing industry, but rather by Business Intelligence (BI). This was further supported by some of the interviewees, as the interviewees were shown an example of a dashboard representing many of the resulting data analytic tools from the market research, whom expressed similar concerns. Furthermore, upon doing a market research in order to find the data analytic tools used by companies within the same industry (manufacturing) has been proven to be challenging. The more traditional data analytic tools, e.g., require a higher level of competence and expertise, whilst the data analytic tool of this context require a "lower level". As one of the purposes behind this context is to identify existing data analytic tools in which can be used by as many people as possible of different backgrounds and competence, a different approach was necessary.

By staying open-minded, it may be possible to use the existing data analytic tools (which is usually perceived as used for business intelligence purposes) to fit the context of this thesis. With this different approach, the benchmarking resulted in 29 different existing data analytic tools. There are more than 29 tools on the market, however the thesis workers deemed that 29 different existing tools was satisfactory to look more into. The 29 different data analytic tools are presented in Table 4.12. with a short description, or comment, of each program. Note that the order of the data analytic tools listed are of no significance, they are purely organized in alphabetical order.

Table 4.12: Results of the list of the existing data analytics tools available in the market.

No.	Program	Description
1	Alteryx	A program that require no coding. Combining data blending, preparation of data, and different types of analytics using the same user interface.
2	Apache Hadoop	An open-source program. This program provides highly distributed, reliable, scalable processing of different and large data sets by only using simple programming.
3	Apache Spark	This program uses programming and is analytics engine for processing large-scale of data.
4	Birst	Connects everything, uses a drag-and-drop layout.
5	Datapine	Good design, drag-and-drop, can make different matrices, sort of simple.
6	Domo	Easy, simple, makes suggestions, helps the user by giving examples.
7	GoodData	An easy program. It uses drag-and-drop system.
8	Google Analytics	Accessible.
9	Google Fusion Tables	Accessible, not that flexible.
10	IBM Cognos (IBM Analytics)	It is more complicated than the other programs but its structure is good and simple to understand.
11	IBM Watson (IBM Analytics)	Ask questions and the program offer different examples and suggestions. Fairly simple to use.
12	InsightSquared	Suggests table, trends but is more sales oriented.
13	KISSmetrics	Fairly complicated to find things and execute, however can create good dashboard.
14	Looker	Build reports that allows the user to explore, drill down for more information and is easy to edit the report. It also uses programming.
15	MATLAB	Requires programming skills, have limited diagrams and graphs.
16	Microsoft Power BI	Simple and easy program that uses drag-and-drop. User-friendly and accessible.
17	Minitab	Only tables, no dashboard, e.g., works like a lower software version of Excel or MATLAB. However, simple to use.
18	Phocas	Nice dashboards, structure, easy simple, works in tables and then diagrams. Is flexible to use.
19	Qlik Sense	Easy to combine datasets from many different data sources. Uses Quick button instead of drag-and-drop when it comes to the datafiles.

Continuation of Table 4.12		
No.	**Program**	**Description**
20	QlikView	Loses in terms of software and layout design, requires many steps to do something simple, no drag-and-drop. Nice dashboards, is easy to use like Word/Excel.
21	RapidMiner	Connecting data, works generally similarly like node-red, no dashboards.
22	SAP Business Intelligence Platform	No dashboard, connecting data in a structural way, not as easy to use.
23	SAP Business Objects Lumira	Good dashboards, easy, works in a structured way, has drag-and-drop.
24	SAP Crystal Reports	Not up to date, creates tables, sort of simple, a lot of steps to do something easy, not really a dashboard. Requires programming.
25	Sisense	A simple and easy program to use when you want to analyze different datafiles and connect them together. The program uses a little bit of programming.
26	Stata	Old software (not up to date), nice diagrams/graphs, not really a dashboard.
27	Tableau	Structured drag-and-drop, nice dashboards, simple, can sort information, make each diagram first and then the dashboards.
28	Yellowfin	It's a simple and easy program that uses programming to make function calls.
29	Zoho Analytics	Is a simple and easy program that uses drag-and-drop for the making of diagram and graphs.

4.3 Result of the Elimination process

The results of the elimination process are presented in this section. All of the 29 different analytic tools are different in many ways, and in order to investigate and verify how well each and every one of them are able to fulfill the requirements specification, the tools were cross referenced with the requirements specification. The elimination process' was executed in three steps and the procedures and results of each step is presented in a chronological order.

During the information collecting regarding each tool and its functions, as the thesis workers were unable to personally verify the legitimacy of the functions through testing the tools, a different approach was necessary. The different approach of verification was conducted through demos and tutorials sourced by the company or other kinds of video-sharing websites, posted by the company themselves or by users. Contacting all the companies of each tool in order to ascertain whether all of the requirements could be fulfilled or not was not feasible within the given time frame. Hence, the thesis workers are aware of how this may have a strong impact on the results for each of the tools. Therefore, the thesis workers strongly recommend proceeding from here on with the aforementioned points in mind, in which may be of importance to the reader of this thesis, upon proceeding with the following results of the elimination process. The full results of the Elimination process: Program vs. Requirement Specification can be found in Appendix F. However, each of the steps are presented in the following sections below.

4.3.1 Elimination process step 1

The first elimination process consists of eliminating the analytic tools that does not fulfill two of the first required criteria stated in the Requirements Specification. The two criteria, criteria 1.1. and 1.2. described in Table 4.13., were fundamental on the process of choosing which programs that would be subjected to further research. Each data analytic tool was first investigated whether they would fulfill criteria 1.1. or not, if it does meet this requirement, the fulfillment of criteria 1.2. was also investigated. If the data analytic tool did not fulfill any of these two criteria, it would be eliminated in the first stage. The two criteria must be fulfilled in order to move on to the next elimination step.

Table 4.13: Criteria 1.1 and 1.2 that must be fulfilled by the data analytic tool found in the Requirements Specification.

Criteria	Description
1.1	Handle many different data and data sources. Data source: e.g. SQL (whereas ERP is an example of a SQL-database). File types: e.g. JSON, CSV
1.2	Uphold certain Technical Standard of 20th century. By a "standard of 20th century", e.g., no Windows XP layout, cannot interact through drag and drop functions and more.

An overview of the results of the data analytic tools from elimination process step 1 is presented in Table 4.14. In the table, the data analytic tools that did not survive the first elimination process are striked-through. The elimination process step 1 resulted in an elimination of 13 programs. The data analytic tools that survived the elimination are the following 16: Birst, Datapine, Domo, GoodData, Google Analytics, IBM Cognos, IBM Watson, Looker, Microsoft Power BI, Phocas, Qlik Sense, SAP Business Objects Lumira, Sisense, Tableau, Yellowfin and Zoho Analytics.

Table 4.14: The results of the data analytic tools from elimination process step 1.

	Program	Criteria fulfilled		Program	Criteria fulfilled
1	~~Alteryx~~	1.1	16	Microsoft Power BI	1.1 and 1.2
2	~~Apache Hadoop~~	1.1	17	~~Minitab~~	1.1
3	~~Apache Spark~~	1.1	18	Phocas	1.1 and 1.2
4	Birst	1.1 and 1.2	19	Qlik Sense	1.1 and 1.2
5	Datapine	1.1 and 1.2	20	~~QlikView~~	1.1
6	Domo	1.1 and 1.2	21	~~RapidMiner~~	1.1
7	GoodData	1.1 and 1.2	22	~~SAP Business Intelligence Platform~~	1.1
8	Google Analytics	1.1 and 1.2	23	SAP Business Objects Lumira	1.1 and 1.2
9	~~Google Fusion Tables~~	none	24	~~SAP Crystal Reports~~	1.1
10	IBM Cognos (IBM Analytics)	1.1 and 1.2	25	Sisense	1.1 and 1.2
11	IBM Watson (IBM ANalytics)	1.1 and 1.2	26	~~Stata~~	1.1
12	~~InsightSquared~~	none	27	Tableau	1.1 and 1.2
13	~~KISSmetrics~~	1.1	28	Yellowfin	1.1 and 1.2
14	Looker	1.1 and 1.2	29	Zoho Analytics	1.1 and 1.2
15	~~MATLAB~~	1.1			

4.3.2 Elimination process step 2

The second step in the elimination process is to eliminate the analytic tools that does not fulfill all the other required (K) criterion described in Table 4.15 below. The tools had to fulfill criteria 2.9 (User-Friendly). In addition to this, if the tools fully fulfilled the required criteria or if they also had some criteria that was left blank (means that more information is required), they would not be eliminated. The reason why they were not eliminated despite not fully fulfilling the required criteria, is because of the way the functions were verified aforementioned. In this elimination process, the end result was that no data analytic tool was eliminated as they all either fully or partially fulfilled the requirement specification. In Table 4.16, representing the results of the data analytics tools of elimination step 2; there were five tools that fully fulfilled the Required criteria (K). These five were the following: Looker, Microsoft Power BI, Phocas, Sisense, and Zoho Analytics. The remaining 11 tools required more information: Birst, Datapine, Domo, GoodData, Google Analytics, IBM Cognos, IBM Watson, Qlik Sense, SAP Business Objects Lumira, Tableau, and Yellowfin. All of these 16 data analytic tools will proceed to elimination process step 3.

Table 4.15: Overview of the required criteria (K)

Criteria	Description	Criteria	Description
1.1	Handle many different data	1.12	Recorded History
1.2	Technical standard of 20th Century	2.1	Simple
1.3	File options	2.2	Clean
1.4	Edit options	2.3	Easy
1.5	View options	2.4	Clarity
1.6	Insert options	2.5	Do-it-yourself
1.7	Format options	2.6	Fast/Quick
1.8	Data options	2.7	Agile
1.9	Solution tools	2.8	Accessible
1.10	Tools	2.9	User-friendly
1.11	Help Search Tool		

Table 4.16: The 16 data analytic tools to be proceeded into elimination process step 3.

Results of elimination process step 2

	Program	Criteria's fulfilled	Unfulfilled	More information required		Program	Criteria's fulfilled	Unfulfilled	More information required
1	Birst	1.1. to 1.11., 2.1. to 2.9.		1.12.	9	Microsoft Power BI	All		
2	Datapine	1.1. to 1.8., 1.10., 2.1. to 2.9.		1.9., 1.11., 1.12.	10	Phocas	All		
3	Domo	1.1. to 1.9., 1.11, 1.12., 2.1. to 2.9.		1.10.	11	Qlik Sense	1.1. to 1.9., 1.11, 1.12., 2.1. to 2.9.		1.10.
4	GoodData	1.1. to 1.11., 2.1. to 2.9.		1.12.	12	SAP Business Objects Lumira	1.1. to 1.8., 1.10 to 1.12. 2.1. to 2.9.		1.9.
5	Google Analytics	1.1. to 1.9., 1.11., 2.1. to 2.9.		1.10., 1.12.	13	Sisense	All		
6	IBM Cognos (IBM Analytics)	1.1. to 1.11., 2.1. to 2.9.		1.12.	14	Tableau	1.1. to 1.11., 2.1. to 2.9.		1.12.
7	IBM Watson (IBM Analytics)	1.1. to 1.7., 1.9. to 1.11., 2.1. to 2.9.		1.8., 1.12.	15	Yellowfin	1.1. to 1.9., 1.11, 1.12., 2.1. to 2.9.		1.10.
8	Looker	All			16	Zoho Analytics	All		

4.3.3 Elimination process step 3

The final step in the elimination process is to proceed with ranking all the analytics tools that survived thus far. For each tool, the amount of X and Y's were summarized and then divided by the total sum of 40x2 = 80. The top 5 programs, marked in blue in table 4.17, are the following: 1. Looker, 2. Yellowfin, 3. GoodData, 4. Microsoft Power BI and 5. Phocas. In first place, Looker scored 97,50 %. The second, third and fourth place is shared by Yellowfin, GoodData, and Microsoft Power BI with a score of 85,00 %. Last but not least, Phocas with a score of 81,25 %. When looking at Table 4.17, the score differences between the top 10 is narrow.

Table 4.17: The results of the data analytic tools with the top 5 marked in blue.

	Software: Data analytic tool	Amount of X	Amount of Y	Total Amount of XY	Percentage
1	Looker	39	39	78	97.50%
2	Yellowfin	37	31	68	85.00%
3	GoodData	34	34	68	85.00%
4	Microsoft Power BI	36	32	68	85.00%
5	Phocas	34	31	65	81.25%
6	Sisense	32	32	64	80.00%
7	Birst	30	33	63	78.75%
8	Zoho Analytics	32	30	62	77.50%
9	Tableau	27	33	60	75.00%
10	SAP Business Objects Lumira	29	28	57	71.25%
11	Domo	28	26	54	67.50%
12	Qlik Sense	28	25	53	66.25%
13	IBM Cognos (IBM Analytics)	22	31	53	66.25%
14	IBM Watson (IBM Analytics)	19	29	48	60.00%
15	Datapine	22	25	47	58.75%
16	Google analytics	23	23	46	57.50%

5
Discussion

As different companies have different needs, the results of data analytic tools suggested in this thesis needs to be adjusted to suit the requirements and needs to the particular company of question. However, the methodology applied in this thesis could be adapted to identify existing data analytic tools and the functional requirements of it. The initial data collected in this thesis, which was collected through the interviews and live mind-map, should be objectively viewed. The reason behind this is due to the initial data having such a great impact on the rest of the methods conducted and depends on the risk of losing important data. In order to minimize the loss of information, and increase the quality of outcome, different steps were taken.

The interviews, transcriptions, keyword identification and live mind-map were all conducted in Swedish before the end-results were further translated into English. This was done in order to stay true to the context of information, and as accurate as possible for as long as possible, to the original source of information. By minimizing the steps of translations required, the translation can be carried out as accurate as possible to the original source of information thus minimizing the risk of losing context during the translation process from one language to another.

In conjunction with this, to counter the risk of misunderstanding the interviewee during the interviews, the interviewers conducted the interviews through conversing and by showing visual illustrations. Notes and audio recordings were also taken off from each interview session. To minimize the risk of missing important information, the reviewing and analyzing process of the collected sets of qualitative data were conducted after each interview session in timely manner. By doing this, it served in improving the questionnaire and as a result of that, also increases the quality of data acquired.

The interviews conducted were semi-structured and the reason as to why this worked well with this context is due to it not following any specific order, but rather by the flow of the session. Hence, as there was a pre-made questionnaire which consists of some basic questions in regard to the context of this thesis serving as a guidance, additional questions could be asked depending on the answers received of which the information could be meaningful.

As the interviewees are of different roles and backgrounds, they all individually held important experience and knowledge regarding the manufacturing industry. However, as there were only 10 interviewees from different departments, the results and its quality could have been further improved by an additional number of attendees from various departments. Especially those of whom worked closely to the production, as it became apparent that these interviewees had a clearer view of what the challenges, problems and needs were. After having conducted the interviews, it became apparent that many of the interviewees did not know what an LDAC was, nor the purpose of it. However, the thesis workers believe that the results and information from this thesis could benefit the case company as it contributes to clarifying and giving a better understanding of what an LDAC is, what it can do and offer suggestions and recommendations.

The results of the thematic analytic methods conducted through a live mind-map should also be taken into consideration, as the keywords identified and categorized could end up looking differently depending on the people behind the creation. However, the creation of the mind-map started off with a blank whiteboard and the reason behind this is to start with a clear mind and to partially or fully eliminate the risk of pre-biased opinions affecting the outcome. At this point the keywords identified from the interviews were already written on post-its. Though, it is worth noting that as there is no definite “right” way on how to identify each keyword correctly. Each of the keywords were captured to the best abilities of the thesis workers with the intention to purely capture the whole picture that each interviewee was trying to depict. Therefore, it is important to recognize the unintended possibility of some keywords perceived differently or overlooked.

With that in mind, some random post-its from the pile of keywords were initially placed randomly on the blank whiteboard. These first random post-its were discussed and placed in different positions or even grouped. Gradually each keyword was placed and in order to increase the quality of the mind-map, the positions of each keyword were questioned and built on arguments through oral discussions held between the thesis workers during the progression. This was done iterative in order to be more critical in the way things are done, as eventually new themes or groups would be formed. This form of execution were done to acknowledge and minimize the possibility of biased opinions each person may have prior to or during the process, in which could affect the results.

As this was done through a live mind-map, a picture was taken of the results as a safety precaution as there, at the time, was no digital version which could be saved for future reference if ever required. In the end, the themes identified are the following: Vision, Challenges, Technologies, LDAC and Required Actions. However, in order to represent the group of keywords more clearly, sub-categories were necessary. The sub-categories identified were the following: SKF, LDAC and possibilities. The results of the mind-map do fall in line with what the thesis workers believe is what the interviewees were trying to depict from the interview sessions.

During the benchmarking process, 29 various data analytic tools were identified which required further investigation. Nevertheless, it is worth noting that there are many more data analytic tools available in the market that might be of interest. However, due to the delimitations of this thesis work it was not feasible to identify and investigate all of them. Hence, in order to fall in line with the given time frame, the 29 data analytic tools were proceeded with further investigation. If the scope of this thesis work were broader it would be possible to further investigate an additional amount of data analytic tools of choice.

The full requirement specification should be viewed objectively as the verification column, which is usually a part of the requirement specification, was left blank. This was the consequences brought forth by the limitations of not having full access to the tools, hence the verification of each tool was not feasible. The different approach taken to verify if the tool offer functions that fulfill the requirements, were through information acquired through demos and tutorials sourced by the company or other video-sharing websites. Hence, it is important to acknowledge that the information may have been biased and therefore could have affected the outcome.

The elimination matrix and process should also be objectively viewed. Upon the creation of the elimination matrix, each of the data analytic tools were cross-referenced with the Requirement Specification criteria. This process was executed separately between the thesis workers and the reason behind this is to take the different characteristics that define individuals into consideration. Therefore, as how every individual work and think do vary, it may lead to different results. Hence, the cross-referencing was conducted separately. The purpose of this execution is to complement with missing information that each thesis worker may have gathered and to be able to work with adequate results.

Furthermore, the investigation of each tool and the source of the information should be taken into consideration. In order to verify the functions of each data analytic tool, the thesis workers were limited to the available information in forms of demos and tutorials, originated from the company themselves or through video-sharing websites. This was due to it not being feasible within the given time frame, wait to gain access of the tools, then test and verify each and every one of them. However, in order to cover for the inconsistencies and limited access, by working separately it created the possibilities of complementing each of the tools with information.

The study conducted for this “Creation of a Local Data Analytics Center” thesis, has been aligned with different aspects that were important to the collaborating company. There are many different data analytic tools and methods available from different sections and industries, with a wide range of different requirements and levels of knowledge and expertise. Compared to the more traditionally focused data analytics utilized by experts (which requires a high level of knowledge), the participants in this study do not have the same level of expertise in the context of data analytics. Expressed differently, the participants of this study have various professional skill sets and knowledge in other areas, e.g., production systems. Hence, the data analytic tools needed to be more user-friendly as the end-users should be able to utilize the software to perform data analytics.

Regardless of this requirement weighing heavily upon the selection of data analytic tools, with an open-mind the thesis workers initiated the investigation by studying all kinds of data analytic tools that is available on the market in order to find tools that possibly could fulfill the criterion stated in the requirements specification. The assessment of existing data analytic tools should be objectively viewed. Additionally, the benchmarking of the tools were limited by some of the further requirements given by the collaborating company. These requirements were 1) the customer base, 2) the reputation, and 3) upcoming trends (software). With this delimitation, it is worth noting that some tools might have been overlooked and important information could have been missed.

With a data analytic tool, it could increase the possibility to share information and be more transparent in what is being presented. In turn, this could increase the credibility and unify the many different opinions into one shared belief during meetings. This tool, could also reduce the amount of time and effort put into searching for the right data information, in creating presentation materials, and the administrative work. Therefore, it makes it possible for the employee to work in a more agile environment where they also can be more productive and effective. Put differently, a data analytic tool could serve in working more effectively in many different areas, in which it is believed could have positive effects when executed correctly. A data analytic tool, implemented into the LDAC, is one of the many ways to provide the means for visualizing, analyzing, predicting, prevent failures, minimize administrative work, unify different beliefs and more which can promote long-term solutions and support conscious decision making.

Despite all the positive possibilities offered by the data analytic tool that are prominent, there are still some additional aspects that should be addressed other than the aforementioned points. One of these aspects is that there may be economic benefits as well as it may serve to reduce the time spent upon meetings, looking through big datasets, or alleviate in preventing manufacturing failures, decision-making and more. However, this point may need further investigation as other areas affected by these changes, e.g., technical support-team, may lead to a shift in where or how much of the investment is going into something. Expressed differently, saving costs in one area does not necessary mean that the company is able to save as much as expected when taking the whole picture into account, as it in turn may create a chain reaction whereas more investments are necessary in other areas.

Another aspect that should be taken into account is how well the LDAC can perform will depend on all the supporting technologies in both software and hardware aspects, as well as the IT architecture and infrastructure. This has become apparent during this study, e.g., due to the lack of a generalized data-hub where all of the different varieties of datasets, from different data sources, currently lack a common connection point. Hence, by looking for the right information from different sources (e.g., data from the various production systems), it decreases the efficiency in terms of work productivity. To enable the creation of insightful data analysis, one must have the option to combine the datasets in different ways depending on what they seek in order to further investigate on.

Furthermore, the results presented by the data analytic tool is very dependent on the legitimacy and quality of the data, what we call Veracity, which may vary. One of the reasons behind this is due to the discrepancies that could occur during the data collection process. Another reason could be in the way the data collection is conducted. In one part of the manufacturing unit, the data collection is automatized with the implemented technologies, whilst in different units it is still done manually filed through paper and pen by different workers each time depending on what shift it could be. The risks this brings is the inconsistency of quality in the datasets, as there is a chance of leaving or missing out on important data, or if the information is not filed correctly or fully, or when the information is passed on to a different employee which could be interpreted differently hence misunderstood. Hence, upon utilizing the data analytic tools, it will still require the user to have a critical mindset, as to not lead the receiver astray with false or incorrect information or make decisions based on false facts.

The inconsistency could also derive from the different levels of analysis performed, which may require the right competencies. In some data analytic tools, upon the creation of the data analysis models, there is a function offered where only the developer or administrator can control the pre-modeling of the data analysis sets. These pre-made data analysis sets could be shared internally, hence used by other employees who can take part of these models by simply input the data of choice. With this function offered, the inconsistency of data analytics could be minimized. Therefore, by implementing some standardization where a limited, skilled and knowledgeable, developers or administrators are the creators of the data analysis models, would decrease the risk of low-quality reports produced. These data analysis-models can be accessible by other users within the corporation, who are able to use them freely but limited to what datasets they wish to be used in these models.

The implementation of a LDAC can lead to some implications to employees, units and organization which may change the ways of working and challenge the mindset of some. Within this context as it is the initial step upon the creation of a LDAC, the organization should put forth strategies to encourage participation and gradual transformation of the data analytics offered. By including people in these strategies, value is added to the project in terms of understanding and trust. Upon the selection of the data analytic tool, it is important to recognize and include the people or team, for whom this tool should be chosen for and especially even more so at this initial stage of LDAC. These people will lay the grounds for how successful the implementation will be, which would either facilitate or hinder the acceptance of this change within the company. These should be taken into consideration, in order to be able to take the opportunities generated by the advantages offered by the data analytic tools.

Upon conducting the study and gone through theory, different methodologies, and ended up with a list of top 16 data analytic tools, the results credibility should be further investigated in terms of testing and verifying. However, the thesis workers were curious as to how well some of the top tools would be able to perform and tried to gain access to the software. However, to no avail. Upon contacting them, or the lack of response thereof, in order to gain access, one must go through the corporation. This was however not feasible to do within the given time frame. Therefore, it was not possible to test the credibility (or rather the functionalities) and how suitable the tool is for each tool listed in the results. Hence, it is recommended that further investigation and testing needs to be done regarding the functionality of the resulted data analytic tools. This thesis investigated existing data analytic tools which could possibly be a part upon the "Creation of a Local Data Analytics Center" and serves to offer prompt solutions and answers to some of the challenges that comes with the technological advancements, hence serving as the first step taken in order to move forward.

6
Conclusions

This thesis evaluated the data analytic tools that are available on the market and what the functional requirements would be for the data analytic tools listed. It serves to add value to the capabilities and possibilities in current practices with a possible solution of working in a standardized way. The results of this thesis were achieved through the conducted literature study, interviews with the stakeholders, benchmarking the available data analytic tools, in which was evaluated in the elimination process. The literature study provided the thesis with information, prior to proceeding with the interviews. The interview study identified 5 different themes; Vision; Challenges; Technologies; LDAC; and Required Action; in which consists of 233 keywords which is the foundation of the 40 identified requirement criteria used in the elimination process. The identified requirement criteria were listed as functional and non-functional criteria set by the stakeholders, in which represents the functional requirements for the data analytic tools available in the market. The results of this thesis should be critically viewed, as this work has only covered a minuscule area of a vast and complex world revolved around big data analytics. To summarize, the answers of the questions will be formulated below.

- What "Data Analytic-tools" are available on the market?

There are various data analytic tools available on the market, whereas some of these data analytic tools were used by all types and sizes of corporations from different sectors and industries. The data analytic tools resulted in 29 different software, which had a customer base, reputation or was an upcoming popular tool. However, the number of tools that fit into this context of work, resulted in 16 different tools. These data analytic tools available on the market are; Birst, Datapine, Domo, GoodData, Google Analytics, IBM Cognos, IBM Watson, Looker, Microsoft Power BI, Phocas, Qlik Sense, SAP Business Objects Lumira, Sisense, Tableau, Yellowfin, and Zoho Analytics. These data analytic tools could possibly be implemented into the initial stage upon the “Creation of a Data Analytics Center”, as in order to progress first requires a step to be taken.

- What are the functional requirements of the data analytic tool?

In this thesis, 40 different functional requirements were identified. These functional requirements for the data analytic tools are the following:

Table 6.1: The 40 identified functional requirements

1	Handle various data files and sources	11	Help search tool	21	Do-it-yourself	31	Different levels of analysis
2	Uphold technical standard	12	Recorded action history	22	Fast/Quick	32	Enable back-tracking
3	File options	13	Data search tool	23	Agile	33	Data input is automatic
4	Edit options	14	Add-Ons	24	Accessible	34	Combine data from different sources
5	View options	15	Filter options	25	User-friendly	35	Combine data from different data types
6	Insert options	16	Access to Data-hub	26	Dynamic/Flexible	36	Use/Identify accurate data
7	Format options	17	Simple	27	Overview	37	Instructions
8	Data options	18	Clean	28	Present results and pinpoint cause	38	Minimize administrative work
9	Solution tools	19	Easy	29	Survey / Overview data	39	Simple report / Share function
10	Tools	20	Clarity	30	Identify and prevent failures	40	Machine learning / AI

Bibliography

Agarwal, R., & Dhar, V. (2014). Editorial-Big Data, Data Science, and Analytics: The Opportunity and Challenge for IS Research. Information Systems Research 25(3):443-448. doi: 10.1287/isre.2014.0546

Ahuett-Garza, H., & Kurfess, T. (2018). A brief discussion on the trends of habilitating technologies for Industry 4.0 and Smart manufacturing. Manufacturing Letters Volume 15 Part B:60-63. DOI: 10.1016/j.mfglet.2018.02.011.

Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., & Ayyash, M. (2015). Internet of things: A survey on enabling technologies, protocols, and applications. IEEE communications surveys & tutorials, 17(4), 2347-2376.

Aronson, J. (1995). A pragmatic view of thematic analysis. The qualitative report, 2(1), 1-3.

Atzori, L., Iera, A., & Morabito, G. (2010). The internet of things: A survey. Computer networks, 54(15), 2787-2805.

Babu, M. S. P., & Sastry, S. H. (2014). Big data and predictive analytics in ERP systems for automating decision making process. 2014 IEEE 5th International Conference on Software Engineering and Service Science, Beijing, 259-262. doi: 10.1109/ICSESS.2014.6933558.

Barham, H. (2017). Achieving Competitive Advantage Through Big Data: A Literature Review. 2017 Portland International Conference on Management of Engineering and Technology (PICMET), Portland, OR, 1-7. doi: 10.23919/PICMET.2017.8125459.

Barham, H., Giadedi, A., Gibson, E., Eljayar, A., & Daim, T. (2018). Evaluating Firm's Readiness For Big Data Adoption: A Hierarchical Decision Model. Huntsville: American Society for Engineering Management (ASEM).

Bryman, A., & Bell, E. (2011). Business Research Methods, 3 ed, Oxford: Oxford University Press.

Buhl, H. U., Röglinger, M., & Moster, F. (2013). Big Data: A Fashionable Topic with(out) Sustainable Relevance for Research and Practice?. Business & Information Systems Engineering, 2, 65-69. doi: 10.1007/s11576-013-0350-x.

Bunse, B., Kagermann, H., & Wahlster, W. (2014). Industrie 4.0: Smart Manufacturing for the Future. Germany Trade & Invest. Retrieved from http://www.inovasyon.org/pdf/GTAI.industrie4.0_smart.manufact.for.future.July.2014.pdf. Retrieved on: 180219

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

Creswell, J. W. (2013). Qualitative Inquiry & Research Design: Choosing Among Five Approaches, 3rd ed, Thousand Oaks, California, USA: SAGE Publications, Inc.

Davenport, T. H. (2014). Big Data @ work: Dispelling the Myths, Uncovering the Opportunities. Boston, Massachusetts, USA: Harvard Business School Publishing Corporation.

Fylan, F. (2005). Semi-structured interviewing. A handbook of research methods for clinical and health psychology, 5(2), 65-78.

Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. Future generation computer systems, 29(7),1645-1660.

Hersman, C., & Fowler, K. (2010). Mission-Critical and Safety-Critical Systems Handbook: Chapter 5 – Best Practices in Spacecraft Development. Design and Development for Embedded Applications. 269-460. Doi: 10.1016/B978-0-7506-8567-2.00005-6

Johannesson, H., Persson, J.-G., & Pettersson, D. (2013). Produktutveckling: Effektiva metoder för konstruktion och design. Stockholm:Liber AB.

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. McKinsey Global Institute. Retrieved from https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation. Retrieved on: 260319

Miorandi, D., Sicari, S., De Pellegrini, F., & Chlamtac, I. (2012). Internet of things: Vision, applications and research challenges. Ad hoc networks, 10(7). 1497-1516.

Nuseibeh, B., & Easterbrook, S. (2000). Requirements engineering: a roadmap. Proceedings of the Conference on The Future of Software Engineering (ICSE'00). 35-46. Doi: 10.1145/336512.336523

Patel, K. K., & Patel, S. M. (2016). Internet of things-IOT: definition, characteristics, architecture, enabling technologies, application & future challenges. International journal of engineering science and computing, 6(5).

Projektledning. 2019. Vad är kravspecifikationer inom projektledning. Retrieved from: https://projektledning.com/kravspecifikation/. Retrieved on: 140519

Popkova, E. G., Ragulina, Y. V., & Bogoviz, A. V. (2019). Revolution of the 21st Century. Studies in Systems, Decision and Control 169. 3-4. doi:10.1007/978-3-319-94310-7

Schmidt, R., Möhring, M., Härting, R., Reichstein, C., Neumaier, P., & Jozinovic, P. (2015). Industry 4.0 - Potentials for Creating Smart Products: Empirical Research Results. Business Information Systems. 16-27. DOI:10.1007/978-3-319-19027-3_2

SKF Group AB (2016). A world of reliable rotation: Annual Report 2016. Retrieved from: https://www.skf.com/irassets/afw/files/press/skf/201703078224-1.pdf. Retrieved on: 100219

SKF Group AB. 2017. SKF enters the world of manufacturing using world-class digital technology. Retrieved from: https://www.skf.com/group/news-and-media/news-search/2017-06-12-SKF-enters-the-world-of-manufacturing-using-world-class-digital-technology.html. Retrieved on 150519

SKF Group AB (2019). SKF History. Retrieved from: https://www.skf.com/group/our-company/organization/skf-history/index.html Retrieved on: 100219

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

Spectral Engines Oy (2018). Material sensing in Industry 4.0 [Online Image] . Retrieved from https://www.spectralengines.com/hs-fs/hubfs/Imported_Blog_Media/SE_artikkeli_kuva_SmartIndustry_v01b-1024x538.png?width=1024&height=538&name=SE_artikkeli_kuva_SmartIndustry _v01b-1024x538.png. 100519

Stal, M. (2014). Chapter 7 – Practices of Software Architects in Business and Strategy – An Industry Experience Report. Economics-Driven Software Architecture. 129-156. Doi: 10.1016/B978-0-12-410464-8.00007-6.

Stankovic, J. A. (2014). Research directions for the internet of things. IEEE Internet of Things Journal, 1(1),3-9.

Sutcliffe, A., & Gulliksen, J. (2012). Chapter 18 – User-Centered Requirements Definition. Usability in Government Systems. 285-300. Doi: 10.1016/B978-0-12-391063-9.00050-X

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

Zhou, K., Liu, T., & Zhou, L. (2015). Industry 4.0: Towards future industrial opportunities and challenges. 2015 12th International Conference on Fuzzy Systems and Knowledge

Discovery (FSKD), 2147-2152. DOI:10.1109/FSKD.2015.7382284

A

The Questionnaire

A.1 The Questionnaire

I **Introduction**

A Can we record the interview?

a We will transcribe the interview which will only be used by us internally to give us support when writing our project report.

B The interview will be on the topic about LDAC.

b It will take approximately 40 minutes of your time.

C Do you want to be anonymous?

c Yes: Can we use your professional role/department as identification?

d No: What's your name?

i What is your professional role?

ii What department do you work for?

D Which departments do you usually cooperate with the most?

e Why?

E With your current role, how long have you been working?

F Are you interested in technology in general?

f What type of technology?

g Why?

II **Information**

G Background information about the project h We are Christina and Ebba and we are Mechanical Engineering students from Chalmers. We are current writing our bachelor's thesis here at SKF which will be about the "Creation of a Local Data Analytics Center".

H Scope of the project and of the interview

i We want to investigate and identify the functional requirements and needs for the data analytic tool that potentially will be used in the LDAC and if this tool would be used on a weekly/monthly basis, depending on your needs and requirements. By identifying the needs, you may have, eventually it may expand to different user-groups within the corporation. The information gathered will be summarized and used in the thesis report and presentation.

III **Acceptance**

I Generally, with your role and experience, what is digitalization for you?

J Do you think that there is any solutions and tolls that can contribute to SKF's competitiveness in the market?

j Which ones?

K In your current role, do you use, or have you seen any type of Data Analytics management?

k Which ones/types?

l When is it used?

m What has worked well?

n What has worked less well? / What limitations were there?

o Is there any function you wish existed?

L Show one example of a dashboard ("LDAC") and explain the image.

p Does this example illustrate what you imagined?

iii No: Why not?

iv Yes: Is there any function you wish existed?

M Have you heard of / used LDAC?

q Yes: Can you tell us more aboute what you have heard/used?

N What potential do you see in an LDAC?

O Do you think that it will be hard to use the technology/tool?

   r Yes: Why?

   s No: Do you think SKF has come far in the digitalization? Could be from a data analytic perspective.

P If a LDAC would be implemented tomorrow, what would get you to use it?

Q Describe a short summary of a normal working day

R What do you think is the most important work task that you do daily?

S Is there anything you think that is necessary to do, but that you think is tiresome/annoying? Could be within data, production, anything.

T Many development projects and similar activities usually involve many people. Do you see any challenges when working in bigger groups?

   s Do you think that a LDAC could be used to partially/fully solve some of the challenges?

U Assuming you have the right set of skills to use this data analytic tool, is there any specific function you think would help you on your daily work tasks?

   t What would you want to get out of a LDAC?

   u What type of data would you want to get out?

   v Why?

V Is there anything else you think is important to include, or that is good?

IV **Summary**

V Summarize the interview and then ask: Have we missed anything that you would like us to bring with us / think about on this project?

W Is there anyone you would like us to talk to?

X To make sure we will ask you once more, do you want to be anonymous?

Y As stated earlier, the gathered information through the interview will all be summarized in a report and presentation. If you would like to, you are much welcome to participate in both.

- w Thank you for coming / Thank you for your time.

B

Live Mind-map

Table B.1: The identified keywords

C

Live mind-map: Identified Keywords

C.1 Overview

Table C.1: Overview of all the themes

Results of Live mind-map: Keywords

Theme: Vision — Category: SKF

n	Keywords
1	Want to evolve & improve
2	Less vertical, more horizontal thinking
3	Better data quality
4	Brind together several potencies (pull together downpipes)
5	Reduce measurment time, more efficient
6	Less administration
7	Reduce costs in general
8	Increase quality
9	High level: IT solutions
10	AI
11	Simpler working methods
12	Reduce external costs (services)
13	SKF: link machine & cloude (ongoing)
14	Transparent for the entire SKF
15	SKF Center (ongoing)
16	Industry 4.0 for the hole SKF
17	Gothenburg & SKF will remain in 10 year
18	Better at analysis level

Theme: Vision — Category: LDAC

n	Keywords
1	Organized
2	Uniform picture of meetings
3	Standardize ways of working
4	Standardize 2-3 analytic tools
5	Used by many (receiver)
6	Few creators
7	Work More Horizontally
8	Keep a clear direction forward
9	Make daily decision easier
10	Make the work easier
11	Make things easier (people)
12	Basis for improvement/changes
13	Improvement opportunities
14	Project management/project drive
15	Drive projects forward
16	Present good facts
17	Problem solving
18	Used for alot
19	Prevent/anticipate problems
20	More time for long-term solutions (no emergency)
21	Streamline work

Theme: Challenges — Category: SKF

n	Keywords
1	Integrate & automate
2	That it is not automated
3	Administration (produce slides
4	A lot of administration
5	Requires a lot of work from the beginning, becomes easier with time
6	Development process of LDAC --> Trial & Error
7	Many meetings
8	Very vertical
9	Focus on one area
10	Different competencies give variations/quality of what is visualized
11	The right conditions for digitalization
12	Different priorities
13	Challenging to get everyone on the same course
14	Challenge: consensus in larger group
15	Ability to change
16	Not so developed
17	From manual to digital, complicated interface
18	Work like an engineer with facts
19	[redacted]
20	Better connection to the network among all machines
21	[redacted]
22	Support for SAP system
23	Data collection should continue despite network problems
24	Maybe need a station between station
25	Digital world? There must be a capacity
26	Has good data, but less good att
27	[redacted]
28	[redacted]

Theme: Challenges — Category: LDAC

n	Keywords
1	SAP & LDAC have access to the same
2	Risk: Different data depending on the module (also in SAP)
3	Input data is important (sensitive!)
4	Risk with the quality depending on the input data
5	Easy way to get data down
6	Big data, collects a lot of data from different places --> find connections and abnormalities
7	Combined data from different sources to the LDAC
8	Quality safe
9	Data integrated between different departments
10	Integration of data flow
11	Organize to integrate data
12	Different levels for different skills

Theme: Technologies — Category: Data-Hub

n	Keywords
1	Common data point where everything goes through/is
2	That the needed data is available
3	Build process & ork methods around data collection
4	More real-time and forward direction (managing KPIs)
5	Real-time (maintenance perspective), prevention before
6	Data-hub (connect all the divferent data sources),
7	Compile the searched data (--> be able to do otherwise, ex.

Theme: Technologies — Sub-category: System

N	Keywords
1	SAP
2	[redacted]
3	[redacted]
4	[redacted]
5	[redacted]
6	[redacted]
7	ERP-system
8	[redacted]
9	Visualization of time-based info
10	[redacted]
11	Excel
12	SharePoint
13	Ishiawa diagram
14	Kaizen
15	Six sigma
16	[redacted]
17	[redacted]

Theme: Technologies — Sub-category: Data

N	Keywords
1	[redacted]
2	SAP
3	[redacted]
4	[redacted]
5	NoSQL
6	SQL
7	Master data
8	Structured data
9	Unstructured data
10	Data is available in local disk, saved in access database
11	[redacted]
12	Situation & deviations with production volumes
13	[redacted]
14	[redacted]
15	[redacted]
16	[redacted]
17	[redacted]
18	[redacted]
19	[redacted]

Theme: LDAC — Category: Functions

n	Keywords
1	Easy & fast to use
2	Do-it yourself
3	Easy to learn
4	User-friendly
5	Create visualization
6	Create insights
7	Informative visualization
8	Create business case
9	Be able to get feedback from X numbers of years
10	See discrepancies
11	Deviations analysis in production
12	Status update (reconciliation)
13	Charts with different aspects (machine- & process data, etc.)
14	Graphs & Charts
15	fast & easy to search for information/status
16	Show different types of connectio
17	Simplification
18	Search-function for example: the data your are looking for
19	problem solving tool
20	Click-view
21	Agile (Dynamic)
22	Dynamic
23	Collected tools
24	Scale away the unimportant (simple functions)
25	History of processed problems

Theme: LDAC — Category: Requests

n	Keywords
1	When running different types of bearings --> the number of transports increases (problems increases
2	Be able to shift from day-day & week-months
3	Day to day: production volume, production rate > see
4	How often the machine stops and the cause
5	Simple & accessible
6	Use in factory logistics
7	Show status in World-Class (on many stations ex AGV, bufferts...)
8	Machine learning
9	Combine data from different data sources
10	Static data combined with real-time data
11	Database search (for data)
12	Be able to enter data automatically
13	Easy connection
14	Analysis team
15	Right people with the right competence
16	LDAC -dependent internal
17	KPI (performance measurement)
18	LDAC can raduce administrative workload?
19	More organized
20	Clarity
21	More independent, less dependent on someone else (to get the data)
22	Instructions (LDAC)
23	LDAC: production-related data (machine & process)
24	Get the right data (LDAC)
25	Overview data (machine- , flows-, plant data)
26	Simple reporting features
27	Early indetify deviations before accidents/serious events

Theme: LDAC — Sub-category: Results

N	Keywords	N	Keywords
1	Result tracking	9	Trends
2	Outgoing quality follow-up	10	Common picture
3	Reflection on results/Metric (KPI don't	11	Common entrance > Analyze together
4	The status of the machines	12	Various POV when looking at the data
5	Number of accidents (Without spare	13	Self-taught/simplification
6	Visualize goals and status report	14	1) Collected Data 2)Analysis 3)Visualize
7	Visualize performance	15	OEE in real time (now-now)
8	Prognosis for the future	16	OEE (Overall equipment effectiveness)

Theme: LDAC — Category: Possibilities

n	Keywords
1	Fact-based decisions
2	Prevent defective products
3	Build away recurring problems
4	Work planned and preventive (rather than last minute)
5	Understand connections
6	Everything fits together
7	Better understanding of the result
8	Discuss & create a basis for business case
9	Develop together (common place)
10	Show & discuss
11	Common place
12	Remove screws that were previously stuck
13	Enable cross-references
14	Challenge old ways of working/performances
15	Support for decisions (Support)
16	View, see the same thing --> solv together
17	Get a conncetion to imlemented improvements/activities
18	Buffer control/status
19	Spend time on analysis & implementation
20	See if problems occur at a particular shift team (ex. incompetence, wrong working method & not technology error)
21	Discipline & structured
22	Improvement activities
23	Collceted image presented
24	Important to be to show people
25	Maintain new thoughts and ideas
26	Flexible to adapt to recipients
27	Customize data combination for the right report/data
28	Make analysis models central/local (less advanced)
29	Adapt content to the stakeholder
30	Reach out with potential, to the users
31	Create root cause analysis
32	Identify risks
33	Identify parameters that affect
34	Identify problems
35	For a selected area, be able to look more closely at the matter
36	Visually get "Here we burst"
37	The imagination sets limits in digitization
38	Complete tasks so that it does not cool down over time

Required Action

n	Keywords
1	Data-
2	SKF IT
3	IT-landscape lacks in general on networks & computers
4	Upload data > User analysis via
5	AWS - Amazon Web Services,

C.2 Identified Keywords (1 out of 5)

Table C.2: Theme Challenges with all of the keywords.

Vision			
SKF		LDAC	
No.	Keywords	No.	Keywords
1	Want to evolve & improve	1	Organized
2	Less vertical, more horizontal thinking	2	Uniform picture of meetings
3	Better data quality	3	Standardize ways of working
4	Brind together several potencies (pull together downpipes)	4	Standardize 2-3 analytic tools
		5	Used by many (receiver)
5	Reduce measurment time, more efficient	6	Few creators
6	Less administration	7	Work More Horizontally
7	Reduce costs in general	8	Keep a clear direction forward
8	Increase quality	9	Make daily decision easier
9	High level: IT solutions	10	Make the work easier (manufacturing)
10	AI	11	Make things easier (people)
11	Simpler working methods	12	Basis for improvement/changes
12	Reduce external costs (services)	13	Improvement opportunities
13	SKF: link machine & cloude (ongoing)	14	Project management/project drive
14	Transparent for the entire SKF	15	Drive projects forward
15	SKF Center (ongoing)	16	Present good facts
16	Industry 4.0 for the hole SKF	17	Problem solving
17	Gothenburg & SKF will remain in 10 years	18	Used for alot
18	Better at analysis level	19	Prevent/anticipate problems
		20	More time for long-term solutions (no emergency)
		21	Streamline work

C.3 Identified Keywords (2 out of 5)

Table C.3: Theme Challenges with all of the keywords.

Challenges			
SKF		LDAC	
No.	Keywords	No.	Keywords
1	Integrate & automate	1	SAP & LDAC have access to the same data
2	That it is not automated	2	Risk: Dfferent data depending on the module (also in SAP)
3	Administration (produce slides regularly)	3	Input data is important (sensitive!)
4	A lot of administration	4	Risk with the quality depending on the input data
5	Requires a lot of work from the beginning, becomes easier with time	5	Easy way to get data down
6	Development process of LDAC --> Trial & Error	6	Big data, collects a lot of data from different places --> find connections and abnormalities
7	Many meetings	7	Combined data from different sources to the LDAC
8	Very vertical	8	Quality safe
9	Focus on one area	9	Data integrated between different departments
10	Different compotencies give variations/quality of what is visualized	10	Integration of data flow
11	The right conditions for digitalization	11	Organize to integrate data
12	Different priorities	12	Different levels for different skills
13	Challenging to get everyone on the same course		
14	Challenge: consensus in larger group		
15	Ability to change		
16	Not so developed		
17	From manual to digital, complicated interface		
18	Work like an engineer with facts		
19	[redacted]		
20	Better connection to the network among all machines		
21	[redacted]		
22	Support for SAP system		
23	Data collection should continue despite network problems		
24	Maybe need a station between station		
25	Digital world? There must be a capacity		
26	Has good data, but less good at analyzing		
27	[redacted]		
28	[redacted]		

C.4 Identified Keywords (3 out of 5)

Table C.4: Theme Technologies with all of the keywords.

Technologies	
	Data-Hub
No.	**Keywords**
1	Common data point where everything goes through/is
2	That the needed data is available
3	Build process & ork methods around data collection
4	More real-time and forward direction (managing KPIs)
5	Real-time (maintenance perspective), prevention before
6	Data-hub (connect all the divferent data sources), automated
7	Compile the searched data (--> be able to do otherwise, ex.

Sub-category		**System**		**Data**
	No.	**Keywords**	**No.**	**Keywords**
	1	SAP	1	[redacted]
	2	[redacted]	2	SAP
	3	[redacted]	3	[redacted]
	4	[redacted]	4	[redacted]
	5	[redacted]	5	NoSQL
	6	[redacted]	6	SQL
	7	ERP-system	7	Master data
	8	[redacted]	8	Structured data
	9	Visualization of time-based infc	9	Unstructured data
	10	[redacted]	10	Data is available in local disk, saved in access database
	11	Excel	11	[redacted]
	12	SharePoint	12	Situation & deviations with production volumes
	13	Ishiawa diagram	13	[redacted]
	14	Kaizen	14	[redacted]
	15	Six sigma	15	[redacted]
	16	[redacted]	16	[redacted]
	17	[redacted]	17	[redacted]
			18	[redacted]
			19	[redacted]

C.5 Identified Keywords (4 out of 5)

Table C.5: Theme LDAC with all of the keywords.

LDAC			
	Functions		Requests
No.	Keywords	No.	Keywords
1	Easy & fast to use	1	When running different types of bearings --> the number of transports increases [problems increases
2	Do-it yourself		
3	Easy to learn	2	Be able to shift from day-day & week-months
4	User-friendly	3	Day to day: production volume, production rate > see
5	Create visualization	4	How often the machine stops and the cause
6	Create insights	5	Simple & accessible
7	Informative visualization	6	Use in factory logistics
8	Create business case	7	Show status in World-Class (on many stations ex AGV, bufferts...)
9	Be able to get feedback from X numbers of years	8	Machine learning
10	See discrepancies	9	Combine data from different data sources
11	Deviations analysis in production	10	Static data combined with real-time data
12	Status update (reconciliation)	11	Database search (for data)
13	Charts with different aspects (machine- & process data, etc.)	12	Be able to enter data automatically
		13	Easy connection
14	Graphs & Charts	14	Analysis team
15	fast & easy to search for information/status	15	Right people with the right competence
		16	LDAC -dependent internal
16	Show different types of connectior	17	KPI (performance measurement)
17	Simplification	18	LDAC can raduce administrative workload?
18	Search-function for example: the data your are looking for	19	More organized
		20	Clarity
19	problem solvning tool	21	More independent, less dependent on someone else (to get the data)
20	Click-view		
21	Agile (Dynamic)	22	Instructions (LDAC)
22	Dynamic	23	LDAC: production-related data (machine & process)
23	Collected tools	24	Get the right data (LDAC)
24	Scale away the unimportant (simple functions)	25	Overview data (machine- , flows-, plant data)
		26	Simple reporting features
25	History of processed problems	27	Early indetify deviations before accidents/serious events

		Results		
Sub-category	No.	Keywords	No.	Keywords
	1	Result tracking	9	Trends
	2	Outgoing quality follow-up	10	Common picture
	3	Reflection on results/Metric (KPI don't	11	Common entrance > Analyze together
	4	The status of the mashines	12	Various POV when looking at the data
	5	Number of accidents (Without spare	13	Self-taught/simplification
	6	Visualize goals and status report (days,	14	1) Collected Data 2)Analysis 3)Visualize
	7	Visualize performance	15	OEE in real time (now-now)
	8	Prognosis for the future	16	OEE (Overall equipment effectiveness)

C.6 Identified Keywords (5 out of 5)

Table C.6: Theme LDAC and Required Action with all of the keywords.

Possibilities		Required Action	
No.	Keywords	No.	Keywords
1	Fact-based decisions	1	Data-
2	Prevent defective products	2	SKF IT
3	Build away recurring problems	3	IT-landscape lacks in general on networks & computers
4	Work planned and preventive (rather than last minute)	4	Upload data > User analysis via
5	Understand connections	5	AWS - Amazon Web Services, Cloud
6	Everything fits together		
7	Better understanding of the result		
8	Discuss & create a basis for business case		
9	Develop together (common place)		
10	Show & discuss		
11	Common place		
12	Remove screws that were previously stuck		
13	Enable cross-references		
14	Challenge old ways of working/performances		
15	Support for decisions (Support)		
16	View, see the same thing --> solv together		
17	Get a conncetion to imlemented improvements/activities		
18	Buffer control/status		
19	Spend time on analysis & implementation		
20	See if problems occur at a particular shift team (ex. incompetence, wrong working method & not technology error)		
21	Discipline & structured		
22	Improvement activities		
23	Collceted image presented		
24	Important to be to show people		
25	Maintain new thoughts and ideas		
26	Flexible to adapt to recipients		
27	Customize data combination for the right report/data		
28	Make analysis models central/local (less advanced)		
29	Adapt content to the stakeholder		
30	Reach out with potential, to the users		
31	Create root cause analysis		
32	Identify risks		
33	Identify parameters that affect		
34	Identify problems		
35	For a selected area, be able to look more closely at the matter		
36	Visually get "Here we burst"		
37	The imagination sets limits in digitization		
38	Complete tasks so that it does not cool down over time		

D

Requirement specifications

D.1 Requirement specifications (1 of 3)

Chalmers	Document type	Requirement Specifications					
	Project	Creation of Local Data Analytics Center (Requirement Specifications for the chosen tool)					
Issued by: Christina Phan		*Created:* 2019-04-02	K = Krav / Required				
Ebba Andrén		*Latest modified: 2019-05-14*	Ö = Önskemål / Request				
eria s			*K/Ö*	*Priority: 1-5*	*Verification method*	*Reference (Stakeholder)*	
Function/s							
	The main function is for the chosen tool to be used in data analysis management		K				
1 Tools and options		Notes					
1.1.	Handle many different data like: Database: SQL, Internal, External ... Filetypes: Excel, CSV ...	Comments: alot of different types/amount of data will be analysed here, from different sources	K			Interviewee	
1.2.	Uphold certain Technical Standard of 20th century	E.g. no Windows XP layout, cannot drag and drop and more	K			Interviewee	
1.3.	File options	E.g. New, Open, Save, Save As, Get Data, Import, Export, Publish, Export to PDF, Share (+interactive), Print, Options & Settings, Help Get Started	K			Interviewee	
1.4.	Edit options	E.g. Undo, Redo, Copy, Cut, Paste, Clear format	K			Interviewee	
1.5.	View options	E.g. Click-View, Page View, Gridlines, Snap Objects, Lock objects, Options to chose what data to view	K			Interviewee	
1.6.	Insert options	E.g. Different visualizations of graphs, diagrams and more	K			Interviewee	
1.7.	Format options	E.g.Font type, Font Size, Font Color, conditional formatting	K			Interviewee	
1.8.	Data options	E.g. where to import data from like SQL and more. Or choose what data to be shown, like country, refine the data	K			Interviewee	
1.9.	Solution tools	E.g. suggested solutions, explore more closer to the results shown	K			Interviewee	
1.10.	Tools	E.g. Spell check, function check (equations) like a red bar signaling that the formula is incorrect, add thresholds	K			Interviewee	
1.11.	Help Search Tool	E.g. search for a "how to" do something	K			Interviewee	
1.12.	Recorded History	E.g. recent files, recent changes and more	K			Interviewee	

D.2 Requirement specifications (2 of 3)

1.13.	Data Search Tool	E.g. trace back the data from where it came from, look for a certain data like date, country, product	Ö	1		Interviewee
1.14.	Add-Ons	E.g. Templates, sheets and more	Ö	2		Interviewee
1.15.	Filter option	E.g. filter different types of data/information (for example, between a certain chosen period of time)	Ö	3		Interviewee
1.16.	Access to general Data-Hub	Comments: To potentially combine many types of different data from different sources. However, currently not implemented, hence impossible to do	Ö	2		Interviewee
2 out		**Notes**				
2.1.	Simple	E.g. does not require coding/scripting	K			Interviewee
2.2.	Clean	To enhance the user experience for every user (does not require you to be an expert in the area)	K			Interviewee
2.3.	Easy	E.g. drag and drop visuals, drag and combine different types of data and more	K			Interviewee
2.4.	Clarity/Tydlighet	E.g. Able to see what has been done in the software, connections between the many different data in the same dashboard	K			Interviewee
2.5.	Do-it-yourself	E.g. be able to handle the software themselves to create their	K			Interviewee
2.6.	Fast/Quick	E.g. can create dashboards within a short amount of time, with precision	K			Interviewee
2.7.	Agile (Quick and Easy?)	To support an agile working environment. These include respect, collaboration, improvement and learning cycles, pride in ownership, focus on delivering value, and the ability to adapt to change	K			Interviewee
2.8.	Accessible	To enhance the user experience for every user, accessible software to many users of different departments. (However, not everyone should be able to edit a published dashboard per say as it can lead to complicated situations - such as misinterpretation)	K			Interviewee
2.9.	User Friendly	One of the goals is to have as many users as possible	K			Interviewee
2.10.	Dynamic/Flexible	E.g. shift between day-to-day report and month-to-month, depending on chosen period of time to review	Ö	2		Interviewee
2.11.	Overview	To be able to quickly have an overview what was done, what has been done, what are the results and more	Ö	3		Interviewee

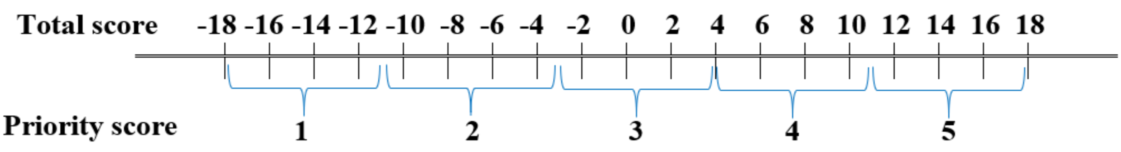
D.3 Requirement specifications (3 of 3)

3	Desired Result	Notes				
3.1.	How often a machine stops and its cause	A requested analysis result from many interviewees	Ö	3		Interviewee
3.2.	Survey/Overview data	A requested analysis result from many interviewees	Ö	3		Interviewee
3.3.	Early-on identify discrepancies prior to failure/serious event	A requested analysis result from many interviewees	Ö	2		Interviewee
3.4.	Different levels of analysis	A requested analysis result from many interviewees	Ö	5		Interviewee
3.5.	Enable back-tracking of X years	A requested analysis result from many interviewees	Ö	4		Interviewee
3.6.	Data input is done automatically	A requested analysis result from many interviewees	Ö	5		Interviewee
3.7.	Combine data from different data sources	A requested analysis result from many interviewees	Ö	5		Interviewee
3.8.	Static data combined with real-time-data	E.g. see trends, A requested analysis result from many interviewees	Ö	5		Interviewee
3.9.	Use/identify accurate data	Comment: Data Analysis is very dependent on the quality of the data you use. E.g. Is able to "clean" data input for you or if you have the option to do so after preview	Ö	5		Interviewee
3.10.	Instructions (LDAC)	E.g. requires standardized instructions (especially at the start-up stage)	Ö	1		Interviewee
3.11.	Minimize administrative workloads	E.g. automatically send a report result of the analysis every friday, from a pre-model	Ö	2		Interviewee
3.12.	Simple report/share functions	E.g. automatically send a report result of the analysis every friday, from a pre-model	Ö	2		Interviewee
3.13.	Machine Learning / AI	E.g. be able to analyse by itself and suggest countermeasure	Ö	5		Interviewee

E

The interface matrix: Priority score of the Request criteria

Interface matrix: Priority score of the request criterias																			
	1.13	1.14	1.15	1.16	2.10	2.11	3.1.	3.2.	3.3.	3.4.	3.5.	3.6.	3.7.	3.8.	3.9.	3.10	3.11	3.12	3.13
1.13.		1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1
1.14.	-1		0	0	0	0	0	0	0	1	0	1	1	1	1	-1	0	0	1
1.15.	-1	0		0	0	0	-1	0	-1	1	0	1	1	1	1	-1	0	0	1
1.16.	-1	0	0		0	1	1	1	0	1	1	1	1	1	1	0	0	0	1
2.10.	-1	0	0	0		0	0	0	0	1	0	1	1	1	1	-1	0	0	1
2.11.	-1	0	0	-1	0		0	0	0	1	0	1	1	1	1	-1	0	-1	1
3.1.	0	0	1	-1	0	0		0	0	1	0	1	1	1	1	-1	-1	-1	1
3.2.	-1	0	0	-1	0	0	0		-1	1	0	1	1	1	1	-1	-1	-1	1
3.3.	0	0	1	0	0	0	0	1		1	0	1	1	1	1	-1	-1	-1	1
3.4.	-1	-1	-1	-1	-1	-1	-1	-1	-1		0	0	0	0	0	-1	-1	-1	0
3.5.	-1	0	0	-1	0	0	0	0	0	0		0	0	0	0	-1	-1	-1	1
3.6.	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0		0	0	0	-1	-1	-1	0
3.7.	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0	0		0	0	-1	-1	-1	0
3.8.	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0	0	0		0	-1	-1	-1	0
3.9.	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0	0	0	0		-1	-1	-1	0
3.10.	-1	1	1	0	1	1	1	1	1	1	1	1	1	1	1		0	0	1
3.11.	-1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0		0	1
3.12.	-1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0		1
3.13.	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	-1	0	0	0	0	-1	-1	-1	
Total score	-16	-4	-2	-9	-4	-2	-3	0	-5	12	4	12	12	12	12	-13	-9	-10	13
Priority score	**1**	**2**	**3**	**2**	**2**	**3**	**3**	**3**	**2**	**5**	**4**	**5**	**5**	**5**	**5**	**1**	**2**	**2**	**5**



F

Elimination matrix: Software vs Requirement Specification

F.1 Overview

Software vs Requirements	1	1	1	1	2	2	2	2	2	0.	1	1	1	1	1	1	2	2	2	2	3	3	3	3	3	0.	2	3	3	3	3	4	4	4	4	4	0.	3	3	3
Alteryx	xy	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Apache Hadoop	xy	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Apache Spark	xy	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Birst	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	x		xy	xy	xy	xy-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		y		xy		y-	xy	xy			y	xy	y-
Datapine	xy	xy	xy	xy	xy	xy	xy	xy		xy				y	y		xy	xy	xy	xy	xy	xy	xy	xy	xy	y	y		y		y			y						
Domo	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy	x	x		xy	xy-	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy		y		xy		y-	x	x		y		x	
GoodData	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy			xy	xy	xy-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy		xy			xy	xy	xy	xy	xy	xy	xy-
Google Analytics	xy	xy	xy	xy	xy	xy	xy	xy	x		xy			xy	xy		xy	xy	xy	xy	xy	xy	y	xy	xy		y		y					xy	x				x	
Google Fusion Tables	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Analytics)	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy			xy	xy		xy	xy	xy	y	xy	xy	y	y	xy	y	xy		y		xy	y	y-		y			y	xy	
Analytics)	xy	xy	xy	xy	xy	xy	xy		xy	xy	xy			xy	xy		y	xy	xy	y	xy	xy	y	y	xy	y	xy		y		y					y	y		xy	y
Insight Squared	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
KISSmetrics	xy	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Looker	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy	xy	xy	xy	xy	rack	xy	xy	xy	xy
MATLAB	xy	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Microsoft Power BI	xy	xy	xy	xy	x	xy	xy	xy	x	xy	x	x	xy	xy	xy	xy-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy		xy	xy	xy	xy	xy		xy	xy	xy	
Minitab	xy	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Phocas	xy	xy+-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy	xy	x-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy		xy-	xy-			x	x	xy	xy	xy	
Qlik Sense	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy	x	x	xy	xy		xy	xy	xy	xy	xy	xy	xy	xy	xy		xy		y		xy			xy	x				x	
QlikView	xy	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RapidMiner	xy	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SAP Business Intelligence Platform	xy	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SAP Business Objects Lumira	xy	xy--	xy	xy	xy	xy	xy	xy		xy	xy	xy		xy	xy		xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy		xy	xy		x				xy		
SAP Crystal Reports	xy	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sisense	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy	xy	xy-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy				xy	xy		xy	xy				xy	xy
Stata	xy	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Tableau	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		y	y	xy		xy	xy	xy	xy	xy	xy	xy	xy	xy	y	y		y		xy	y	xy	xy	xy	x		y	xy	
Yellowfin	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy	x	xy		xy	x-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy		xy	x	xy-	xy	xy	x	x	xy	xy	xy-
Zoho Analytics	xy	xy	xy	xy	xy	xy	xy	xy	x	xy	xy	x	x	xy	xy	xy-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		y		xy	xy		xy	xy				xy	

(red)	- Required (K) must be fulfilled in elimination process step 1
(grey)	- Required (K) must be fulfilled in elimination process step 2
	= Request (Ö) criterias
empty	- more information required
-	- Does not fulfill / has not been looked on
x / y (-)	- Does fulfill partially (weakly)
x / y (+)	- Does fulfill partially (strongly)

F.2 Elimination matrix: Software vs Requirement Specification (1 of 2)

Software vs Requirements Specification Criterias	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10.	1.1	1.1	1.1	1.1
Alteryx	xy	-	-	-	-	-	-	-	-	-	-	-	-	-
Apache Hadoop	xy	-	-	-	-	-	-	-	-	-	-	-	-	-
Apache Spark	xy	-	-	-	-	-	-	-	-	-	-	-	-	-
Birst	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	x		xy	xy
Datapine	xy	xy	xy	xy	xy	xy	xy	xy		xy				y
Domo	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy	x	x	
GoodData	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy			xy
Google Analytics	xy	xy	xy	xy	xy	xy	xy	xy	x		xy			xy
Google Fusion Tables	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IBM Cognos (IBM Analytics)	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy			xy
IBM Watson (IBM Analytics)	xy	xy	xy	xy	xy	xy	xy		xy	xy	xy			xy
Insight Squared	-	-	-	-	-	-	-	-	-	-	-	-	-	-
KISSmetrics	xy	-	-	-	-	-	-	-	-	-	-	-	-	-
Looker	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy
MATLAB	xy	-	-	-	-	-	-	-	-	-	-	-	-	-
Microsoft Power BI	xy	xy	xy	xy	x	xy	xy	xy	x	xy	x	x	xy	xy
Minitab	xy	-	-	-	-	-	-	-	-	-	-	-	-	-
Phocas	xy	xy+-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy
Qlik Sense	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy	x	x	xy
QlikView	xy	-	-	-	-	-	-	-	-	-	-	-	-	-
RapidMiner	xy	-	-	-	-	-	-	-	-	-	-	-	-	-
Platform	xy	-	-	-	-	-	-	-	-	-	-	-	-	-
Lumira	xy	xy--	xy	xy	xy	xy	xy	xy		xy	xy	xy		xy
SAP Crystal Reports	xy	-	-	-	-	-	-	-	-	-	-	-	-	-
Sisense	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy
Stata	xy	-	-	-	-	-	-	-	-	-	-	-	-	-
Tableau	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		y	y
Yellowfin	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy	x	xy	
Zoho Analytics	xy	xy	xy	xy	xy	xy	xy	xy	x	xy	xy	x	x	xy

(red)	= Required (K) must be fulfilled in elimination process step 1
(grey)	= Required (K) must be fulfilled in elimination process step 2
	= Request (Ö) criterias
empty	= more information required
-	= Does not fulfill / has not been looked on
x / y (-)	= Does fulfill partially (weakly)
x / y (+)	= Does fulfill partially (strongly)

F.3 Elimination matrix: Software vs Requirement Specification (2 of 2)

1.2	1.2	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	2.10.	2.1	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	3.10.	3.1	3.1	3.1
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
xy	xy-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		y		xy		y-	xy	xy			y	xy	y-
y		xy	xy	xy	xy	xy	xy	xy	xy	xy	y	y		y		y			y						
xy	xy-	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy		y		xy		y-	x	x		y		x	
xy	xy-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy		xy			xy	xy	xy	xy	xy	xy	xy-
xy		xy	xy	xy	xy	xy	xy	y	xy	xy		y		y					xy	x				x	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
xy		xy	xy	xy	y	xy	xy	y	y	xy	y	xy		y		xy	y	y-		y			y	xy	
xy		y	xy	xy	y	xy	xy	y	y	xy	y	xy		y		y					y	y		xy	y
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
xy	xy-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy	xy	xy	xy	xy	ack	xy	xy	xy	xy
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
xy	xy-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy		xy	xy	xy	xy	xy		xy	xy	xy	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
xy	x-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy		xy-	xy-			x	x	xy	xy	xy	
xy		xy	xy	xy	xy	xy	xy	xy	xy	xy		xy		y		xy			xy	x				x	
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
xy		xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy		xy	xy		x				xy		
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
xy	xy-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy				xy	xy		xy	xy				xy	xy
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
xy		xy	xy	xy	xy	xy	xy	xy	xy	xy	y	y		y		xy	y	xy	xy	xy	x		y	xy	
xy	x-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		xy		xy	x	xy-	xy	xy	x	x	xy	xy	xy-
xy	xy-	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy	xy		y		xy	xy		xy	xy				xy	