LINKÖPING UNIVERSITY | IEI
30 HP MASTER THESIS
INDUSTRIAL ENGINEERING AND MANAGEMENT | QUALITY MANAGEMENT
SPRING TERM 2020 | LIU-IEI-TEK-A--20/03736--SE

DIGITAL TOOLS FOR ANALYSIS AND IMPROVEMENT

- INCORPORATION OF DIGITAL TOOLS IN IMPROVEMENT PROCESSES

CARL MAGNUSSON
ESKIL NORDLUND

PROPIA SUPERVISOR: MARTIN ERICSSON
LIU SUPERVISOR: PETER CRONEMYR
LIU EXAMINER: MATTIAS ELG

Preface

This master thesis constitutes the final work of five years of studies at Linköping University. The thesis is written within the subject of Quality Management by the two master's students Carl Magnusson and Eskil Nordlund in partnership with the thesis constituent Propia AB.

We would like to express our deepest gratitude towards Martin Ericsson, our supervisor at Propia and Peter Cronemyr, our supervisor at Linköping University. Their continuous support and guidance have been invaluable during the course of this thesis. We would also like to thank the other consultants at Propia for their warm welcome and inclusiveness from the first day.

In addition, we would like to thank all respondents from Tekniska Verken, TitanX, Ericsson, and Sandvik for their participation and for enabling this study by sharing their experiences and knowledge. Finally, we would like to thank our examiner Mattias Elg and opponents Marcus Eriksson and Maria Beck-Friis Lundholm for rewarding seminars and discussions.

Carl Magnusson and Eskil Nordlund
Linköping, May 2020

Abstract

This thesis was conducted on behalf of the consultancy firm Propia AB based in Norrköping, Sweden with the goal to develop a method for analysis and improvement that incorporates digital tools and techniques. *Digital tools* have in this study been defined as “Tools characterized by electronic and especially computerized technologies”.

The study was carried out by abductive approach using both literature studies and interviews. The literature focused on improvement work and how digital tools can be incorporated which resulted in a theoretical framework based on Six Sigma’s DMAIC structure with incorporated digital tools. The tools were ranked according to *difficulty of use* and *added value* and matched against maturity levels for process and digital maturity to create an indication for businesses what tools to use. The output of the theoretical analysis was then used as a basis for the empirical studies.

Empirical studies were conducted in the form of semi-structured interviews with respondents from five different companies from different industries. Input was provided on improvement work, tools used in their improvement work and digital and process maturity and the interaction between the areas. Through combining the empirical study with previous theory, the theoretical framework was revised from a few key takeaways.

The thesis’ result is presented in the form of two improvement structures, one for larger problem solving and improvement opportunities and a smaller continuous improvements structure. The frameworks are presented as flow charts with yes and no questions to guide through the improvement process. Digital tools are then matched to each step in the improvement process and examples of use areas are presented. The individual digital tools are also connected to maturity levels to customize the improvement process to any business.

Sammanfattning

Denna masteruppsats utfördes på uppdrag av konsultfirman Propia AB, baserad i Norrköping, Sverige med målet att utveckla en strukturerad metod för analys och förbättring som integrerar digitala verktyg och metoder. Digitala verktyg har i denna studie definierats som *verktyg som karakteriseras av elektroniska och framför allt datoriserad teknologi*.

Studien utfördes med en abduktiv ansats genom en litteraturstudie och intervjuer. Litteraturstudien innefattade främst förbättringsarbete och hur digitala verktyg har integrerats i tidigare studier. Genom inledande teoretisk analys togs ett ramverk fram baserat på Six Sigmas DMAIC-struktur med digitala verktyg kopplat till de olika faserna. De digitala verktygen rankades därefter utifrån *svårighet* och *värdeskapande* och matchades mot nivåer för processmognad och digital mognad. Ramverket tillsammans med mognadsnivåerna agerade sedan som underlag för den empiriska studien.

Den empiriska studien genomfördes i form av semistrukturerade intervjuer med respondenter ifrån fem olika företag verksamma inom olika branscher. Intervjuerna fokuserade på områdena förbättringsarbete, verktyg som företagen använder i sitt förbättringsarbete samt processmognad och digital mognad och samspelet däremellan. Genom att kombinera den empiriska studien med tidigare teori reviderades det teoretiska ramverket till ett bredare och empiriskt förankrat ramverk.

Uppsatsens resultat presenteras i form av två förbättringsstrukturer. En för större problem och förbättringsmöjligheter och en för mindre ständiga förbättringar. Ramverken presenteras som flödesscheman med ja och nej frågor vilka vägleder användaren genom förbättringsprocessen. Till varje steg finns adekvata digitala verktyg som kan integreras i förbättringsprocessen med exempel på förbättringsområden. De enskilda digitala verktygen kopplas också till mognadsnivåer för att kunna anpassa modellerna till företag oberoende mognadsnivå.

Content

List of Figures

List of Tables

1 Introduction

In the introduction chapter background to the problem, purpose, research questions and delimitations are presented.

1.1 Background

Bernard Marr (2018) writes that 90 % of the data in the world have been created in the last two years. The increasingly digitalized world provides huge amounts of data which opens opportunities to work smarter and more efficient within business processes. Over the last decade the market has been flooded with software to visualize and analyze the data generated. Yet, many organizations have a hard time exploiting that information efficiently in their operations (Propia AB, 2019). *Digital tools* are a broad term which is defined differently depending on who you ask. Steils and Hanine (2019) describes the term as “Tools characterized by electronic and especially computerized technologies” which will also be the definition used in this thesis. Examples of digital tools includes everything from mobile phones, Internet of Things, sensors, statistical software, and programming techniques such as data mining and machine learning.

Lee, Kao and Yang (2014) describes an emerging challenge for industrial companies as the amount of data to analyze continues to increase. The industrial companies’ manufacturing systems are not ready to handle it as they are intended for lesser amount of data. However, changing system is often costly resulting in keeping the old systems for long. Furthermore, the enormous amounts of data also bring disadvantages when trying to analyze it. Correlations can be found due to the sheer amount of data even though it has no relation. Calude and Longo (2016) write “paradoxically, the more information we have, the more difficult is to extract meaning from it. Too much information tends to behave like very little information.” Propia’s experience is that there is a notion that simply visualizing data will automatically lead to analysis and improvement. However, this is not always the case, on the contrary solely visualization might instead result in wrongful and unnecessary conclusions and decisions. (Propia AB, 2019)

Propia AB is a consultancy firm based in Norrköping, specializing in process management, change management and business development (Propia AB, 2019). Propia has identified that there is a gap when it comes to structured methodology for cause and effect analysis connected to the digital tools. They have observed that companies manage to visualize the data but fail to analyze it in a correct way, which leads to wrongful conclusions. Propia see a potential to use root cause analysis such as Six Sigma in connection with digital tools and techniques in the different phases, Define-Measure-Analyze-Improve-Control (DMAIC) in order to help companies to draw the correct conclusions. Gupta, Modgil & Gunasekaran (2020) argues that new big data applications within Lean Six Sigma projects can pay off in tangible and intangible benefits to firms and Laux, et al. (2017) concludes that Six Sigma and big data combined can mitigate the separate weaknesses and result in an integrated improvement system.

1.2 Purpose and Research Questions

Digital tools to support improvement work exist in wide variety, however a structured method to showcase when and how to utilize them is lacking. In regard to this problem this master thesis’ first aim is to present a theoretical framework with existing tools and methods incorporated into the DMAIC-cycle to highlight gaps and provide directions when to utilize what tool or technique.

An exploration of how and in what context digital tools are being used today, as well as previous efforts to combine digital tools and improvement processes will be investigated. The second aim of the thesis is to establish a framework that considers process and digital maturity when integrating digital tools and techniques into different improvement processes. This framework is then meant to act as a guide for businesses to implement structured improvement work with digital tools according to their digital maturity.

The research questions that will be answered to support the thesis' aim is:

Research Question 1

What steps of the data analysis and improvement process lack or have limited support by the digital tools of today?

Research Question 2

How can today's digital tools be used in a structured analysis and improvement process?

Research Question 3

How can a structured analysis and improvement process be adjusted to match a business' digital maturity?

1.3 Delimitations

The framework will not consider back end data infrastructure and architecture which includes servers and data layers that need to be in place to use some of the tools and techniques. The back-end infrastructure is unique for all businesses and difficult to generalize. Furthermore, the study was performed within quality management and emphasis was therefore focused on the abilities of the tools. Systems used in day-to day operations such as Enterprise Resource Planning and Manufacturing Executions Systems will be neglected from the term digital tools. No implementation or testing of the framework will be included in this thesis.

2 Methodology

The following chapter details the reasoning and selection of research methods and scientific approach used in order to answer the thesis' research questions. A qualitative method with an abductive approach was chosen to conduct this study and both primary data in the form of semi-structured interviews and secondary data in the form of a literature study has been used.

2.1 Research Design

The aim of this study was to develop two frameworks. The first for showcasing the gap between the structured improvement process DMAIC and digital tools of today. The second acting as a guide for integrating digital tools in structured improvement process according to a business digital maturity. The first framework was purely theoretically based whereas the second one was a combination of theoretical foundation and empirical validation. The start of the thesis was performed by conducting a pre-study to get an overview of the subject. This was followed by a systematic literature study on relevant topics. Based on the literature study a theoretical analysis was performed which resulted in an initial framework. The framework was then used as a base for the empirical study and further analysis. After the empirical study, the final framework emerged after revision according to the empirical findings in the analysis phase. The process can be seen in Figure 1.

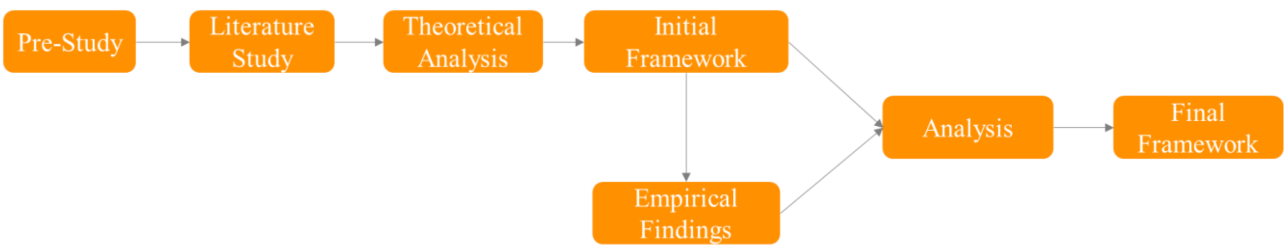

Figure 1. Design of the study.

There are four types of designs that can be used when performing a study, these are exploratory, descriptive, explanatory, and prescriptive. *Exploratory* design is used when knowledge within an area are limited and the goal is to achieve a basic understanding. *Descriptive* design is used when the goal of a study is to describe relations within an area where there is basic knowledge and understanding. *Explanatory* studies seek to both describe and explain and are used to gain deeper knowledge and understanding. *Prescriptive* studies are used when there are existing knowledge and understanding of a research area and the goal is to provide guidance and recommendations. (Björklund & Paulsson, 2012). This study utilized an exploratory design as the goal of the study is to explore the gap between existing structured improvement methods and digital tools and present a framework which aims to help business with structured improvement work.

Jørgensen & Rienecker (2014) argues that the research questions in a well-structured report should have a strong relation to each other. Therefore, the research questions in this thesis are structured in a chronological order where the previous one lays the foundation for the next. The answers to the three research questions in this thesis will all be based on theoretical findings with the last two being validated with empirical findings. A visualization of the approach can be seen in Figure 2.

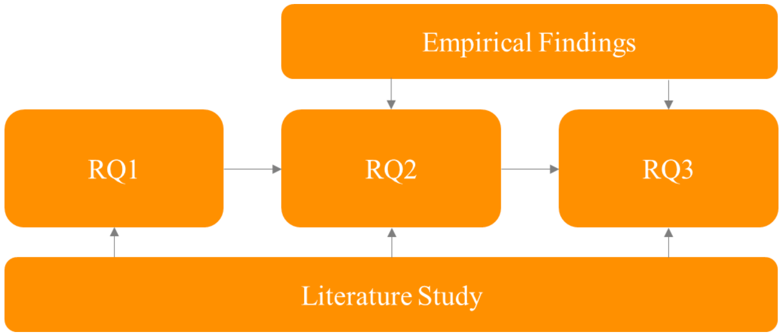

Figure 2. Approach to research questions

2.2 Research Method

There are three common scientific approaches to conduct a study, *Induction*, *Deduction* and *Abduction*. An inductive approach starts with empirical findings to try and build theory. Deduction instead starts with existing theories and make predictions which are then attempted to verify through empirical findings. Abduction is an interaction of deduction and induction where an iterative process between theoretical and empirical findings is used to develop new knowledge based on existing theories. (Björklund & Paulsson, 2012). An abductive approach facilitates the development of existing theory (Saunders, et al., 2009). The framework presented in the report is based on theoretical findings and developed using empirical findings together with the frame of reference and theoretical analysis, which is why this study is of abductive nature.

2.3 Pre-study

In the preface of the study multiple semi structured, broader literature searches were conducted to familiarize with the subjects of the study. Gradually, as the understanding of the subject areas increased more targeted keywords could be used in the more extensive literature study. The primary topics in the initial searches were Six Sigma, Business Intelligence (BI) and big data. Unstructured interviews at the case company Tekniska Verken were held to gain an introductory view of the organization and their current challenges to gain a holistic view of the subject as well as a practical introduction to a BI software.

2.4 Literature Study

Literature studies are used to gather secondary data, which mean that it was gathered with a different purpose than for the study in question. It is a method that enable gathering of large amounts of information in short time and with almost no financial resources. However, as it is secondary data, it is important to be aware that the information might be biased or incomprehensive. (Björklund & Paulsson, 2012). A literature study was initiated after the pre-study was finished to gain further insight into the relevant subject areas. The literature used consisted of articles and books found through Linköping University online library UniSearch as well as the databases Scopus, Web of science and Google Scholar. Most sources used have been peer-reviewed articles as they have been audited by other authors and can be seen as more trustworthy (Alexanderson, 2016). Publications were read and sorted based on title, abstract and full text to make the literature study as efficient as possible. The sorting of articles started with reading the abstracts and if it was deemed relevant to the study the full article was read, otherwise it was discarded. This approach comes with the risk of discarding articles that contain relevant information not mentioned in the abstract, however, due to time limitations it was the approach used. Furthermore, several of the articles that were deemed relevant were later rejected as the content in the article showed to be out of the scope for the study.

To ensure the quality, several sources were used to complement each other. Björklund & Paulsson (2012) states that two or more separate sources presenting similar information should be regarded as more reliable. As the study included topics are relatively modern subjects of research, the date of the literature was of importance. The fast-paced development renders research outdated fast, which is why the date of the articles was also taken into consideration.

2.5 Empirical Study

There are two types of empirical studies, quantitative and qualitative studies. Quantitative studies are performed when handling numerical data that can be measured. Often times these studies include larger data sets or surveys which can be analyzed by statistical calculations and software. The drawback with quantitative studies is that everything cannot be measured. Qualitative studies are instead used to attain a deeper understanding of a subject or problem (Björklund & Paulsson, 2012). Pajo (2017) describes it as bringing in various perspectives in an attempt to enrich the picture when looking at an issue. This is often achieved by interviews or observations which enables a more flexible approach. The drawback is that the result is less generalizable (Björklund & Paulsson, 2012). This study utilized a qualitative approach in the form of interviews to gather data as the aim was to get a deeper understanding of the subject. Furthermore, the limited number of participants in the study may have resulted in a skewed result if a quantitative approach was used which appeal to the use of qualitative approach.

When conducting interviews there are three main types that can be used, unstructured, semi-structured or structured. The differences are how strictly a series of questions worked out ahead of time is followed (Pajo, 2017). In structured interviews, all questions are predetermined and handled in a specific order whereas unstructured interviews may not have actual questions prepared (Pajo, 2017). The semi-constructed method is used by standardizing questions and asking the same questions to all respondents; however, follow-up questions can differ depending on the answer given by the respondent (Björklund & Paulsson, 2012). The method more than often resembles a conversation than a strict interview where the respondent can feel more comfortable and is able to give more nuanced answers on the topic.

The interviews in this thesis was conducted by using a semi-structured approach as it encourages discussion around the pre-determined questions. This was selected as the study uses an exploratory design aimed to explore the subject further and the diversity of interviewees was another aspect of the choice. Respondents were from different companies and had diverse positions and knowledge and with the semi-structured format they could elaborate further on their particular methods and obstacles faced. Due to this a semi-structured interview worked well as the interviewees received the same questions but were encouraged to discuss their answers. The main purpose with the empirical study was to get insight into obstacles experienced in improvement work and the usage of improvement structure and digital tools. This was a way to pinpoint in what parts of improvement work often fall short.

All the interviews were performed digitally with both authors present. One of the authors was designated to take notes during the interviews while the other one lead the interviews and asked the questions. The interviewee was informed of the purpose of the study and interview before it started. They were also informed that they had the possibility of cancelling the interview at any time and refrain from answering any question. Respondents were also asked permission for the interview to be recorded, all interviews were thereafter transcribed and examined for keywords in order to thematize them before summarizing them. A summary of the interviews can be seen in Table 1.

Table 1. Overview of interviews held

	Date	Time (h)	Present	Company	Recorded
Interview 1 (Unstructured)	2020-02-13	1,5	Two consultant, Business developer	Propia & Tekniska Verken (BBE)	No
Interview 2 (Unstructured)	2020-02-20	1,5	Consultant, Department Manager Business Development	Propia & Tekniska Verken (NÄT)	No
Interview 3 (Unstructured)	2020-02-24	1,5	Consultant, Business Developer, BI-responsible	Propia & Tekniska Verken (Bixia)	No
Interview 4 (Semi-structured)	2020-03-26	1	Quality Manager	TitanX	Yes
Interview 5 (Semi-structured)	2020-04-07	1	Business Developer	Tekniska Verken (BBE)	Yes
Interview 6 (Semi-structured)	2020-04-08	1	Business Developer	Tekniska Verken (BBE)	Yes
Interview 7 (Semi-structured)	2020-04-08	1	Business Developer, Department Manager Business Development and Support	Tekniska Verken (NÄT)	Yes
Interview 8 (Semi-structured)	2020-04-14	1	Department Manager Business Development	Tekniska Verken (BBE)	Yes
Interview 9 (Semi-structured)	2020-04-16	1	Business Developer	Tekniska Verken (BBE)	Yes
Interview 10 (Semi-structured)	2020-04-21	1	Continuous Improvement	Sandvik	Yes
Discussion seminar	2020-04-23	1	Consultants	Propia	No
Interview 11 (Semi-structured)	2020-04-24	1	Senior process manager, Business performance & improvement manager	Ericsson	Yes

2.6 Validity, Reliability and Objectivity

Validity, reliability, and objectivity should be taken in account to increase the credibility of a scientific study. All studies should strive to maximize these three, however, they need to be weighed against necessary resources, such as time. Validity can be described as to what extent what is intended to be measured is in fact measured. Reliability is consistency in the measurement, to what degree the same values are obtained if the investigation is repeated. Objectivity is the extent the study is affected by external values, and authors bias. (Björklund & Paulsson, 2012). Validity and reliability are somewhat connected, if reliability is missing, so is validity. However, reliability does not imply validity as consistency can be present, but still measure the wrong thing (Bell, 2016).

To increase validity and reliability different versions of triangulation, i.e. including multiple perspectives can be used. It can be performed by using multiple sources, theories and persons evaluating used material (Björklund & Paulsson, 2012). Objectivity of the study can be increased by motivating and clearly state choices made during the study. When using what others have written, the content of the source should be stated as objectively as possible by communicating correct facts and not only present facts supporting the own view (Björklund & Paulsson, 2012). Recognizing and being aware of personal biases can also help in getting closer to objectivity (Pajo, 2017).

The aim of this master thesis was to present a framework that can be used by Propia and be applicable to several different businesses, therefore a high validity, reliability and objectivity was sought after. Validity and reliability of the thesis was increased by the use of data triangulation, the method of using several different sources to validate a text. The sources were chosen with caution and primarily from books and peer reviewed articles and the assertions made in the sources was then supported with different sources in order to increase validity. Objectivity was increased by using objective and non-emotive language through the report to avoid authors bias. Furthermore, different perspectives were presented in order to give the reader a comprehensive depiction of the literature and not a distorted representation which fits the authors narrative. Lastly the generalizability was increased by interviewing several different companies in the empirical study. This increases the different perspectives and give a more nuanced framework which is applicable to more businesses.

2.7 Ethics

The Swedish Research Council describes two aspects of ethics, the nature of the research and the researchers conduct (Swedish Research Council, 2019). The nature of the research includes being impartial and unbiased towards both personal and outside interests. The Swedish Research Council has compiled a list of eight rules which summarizes a good ethical oath.

1. "You shall tell the truth about your research.
2. You shall consciously review and report the basic premises of your studies.
3. You shall openly account for your methods and results.
4. You shall openly account for your commercial interests and other associations.
5. You shall not make unauthorised use of the research results of others.
6. You shall keep your research organised, for example through documentation and filing.
7. You shall strive to conduct your research without doing harm to people, animals or the environment.
8. You shall be fair in your judgement of others' research." (Swedish Research Council, 2017).

ALLEA (All European Academies) published a guideline report “The European Code of Conduct for Research Integrity” in which they present four cornerstones for good ethical practices. These four foundations are, *Reliability*, *Honesty*, *Respect* and *Accountability* (ALLEA - All European Academies, 2017). These four can be condensed to ensuring an unbiased research with high quality and authenticity which has been conducted with respect towards colleagues, society and the environment as well as being published in order for others to learn and build on (ALLEA - All European Academies, 2017).

In order to conduct interviews that fulfill the rules and foundations mentioned above several measures was taken into consideration. Before every interview, an explanation of the purpose and the aim of the study was given. This was done to ensure that the interviewees were aware and comfortable with their participation. Furthermore, the interviewees were also informed that they could cancel the interview at any time, their names would be anonymized and would only be recorded if their permission were granted.

3 Frame of Reference

The frame of reference present relevant theories that act as a basis to answer the research questions. It is divided according to the areas of improvement work and processes, digital tools and technologies and how the areas can interact.

3.1 Total Quality Management and Improvement Processes

Quality is commonly defined as the ability of a product or service to satisfy or exceed the customer's expectations. Customer focus is central in quality management and the customer can be defined as the one that value is created for. (Bergman & Klefsjö, 2010)

Total Quality Management (TQM) is described by Bergman & Klefsjö (2010) as striving to fulfill and preferably exceed customer's needs and expectations with minimal resource consumption through continuous improvement work where everyone is involved and has focus on the organization's processes. *Committed leadership* must be practiced on all levels of the organization to create a culture for successful and sustainable improvements. *Focus on customers* implies finding out what the customer wants and needs, then try to fulfill and exceed those needs and expectations. Facilitating opportunities for *everyone to be committed* and participate in the improvement work is an important mean to achieve quality improvement as it increases responsibility as well as professional pride. A process is chain of repeated activities performed to create value for a customer, and a company can be described as a network of processes. *Focusing on processes* means to redirect attention from products to the chains of activity producing them and asking how the results are produced. The rule of *continuous improvements* is that there is always a way to get improved quality using less resources, and anyone who stops improving stops being good. By analyzing data and separate correct from irrelevant information, knowledge should be extracted and serve as basis for decisions, i.e. *base decisions on facts*. Figure 3 show how these concepts act as cornerstones for TQM. (Bergman & Klefsjö, 2010).

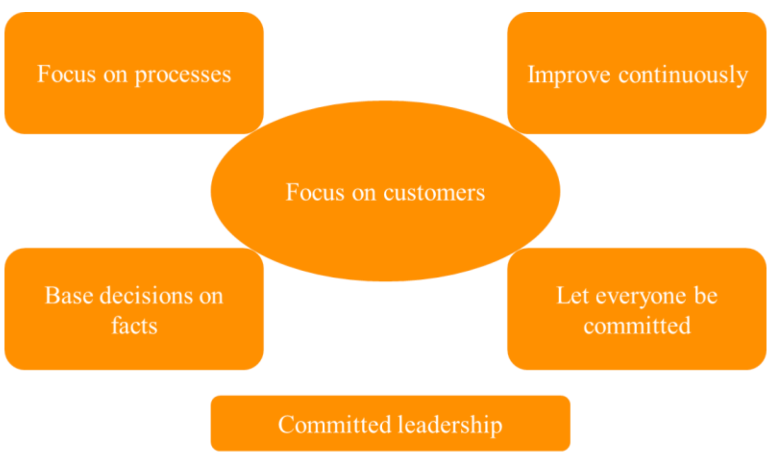

Figure 3. Cornerstones of Total Quality Management (Bergman & Klefsjö, 2010)

The cornerstone model incorporates both intangibles as developing an organizational culture with customer focus, committed leadership, and including everyone as well as tangibles with improving continuously by basing decisions on facts and focusing on processes. Intangibles, also known as soft values, are often difficult to measure and includes company culture, values and commitment which are vital when implementing and working with TQM. Tangibles, or hard values, on the other hand are easier to measure. This includes economic results, measurements, savings etc. This thesis will focus on the three tangible cornerstones' "Focus on processes", "Improve continuously", and "Base decision on facts" as the thesis purpose is to integrate improvement processes with digital tools and digital maturity.

3.1.1 Methodologies in Improvement Processes

There exist a variety of different improvement processes such as PDCA, A3, 8D and DMAIC (Six Sigma). The aim of all these models are continuous improvement which raises the question, what is the difference between them? Liesener (2013) has composed a figure which showcases the different phases in the models compared to each other which can be seen in Figure 4. He argues that all four models utilize a scientific way of tackling a problem in order to solve it, i.e. *basing decisions on facts* through statistical analysis. The objective of all improvement processes is to find and eliminate root causes to problems in order to ensure better quality at a reduced cost, in line with the cornerstone *improve continuously*.

PDCA	DMAIC	A3	8D/PSP
Plan	DEFINE	Clarify the Problem	1. Create Team & collect Information
	MEASURE	Break down the Problem	2. Describe the Problem
		Set a Target	3. Define Containment Actions
	ANALYZE	Analyze the Root Cause	4. Analyze the Root Cause
		Develop Countermeasures	5. Define possible corrective Actions
Do	IMPROVE	See Countermeasures	6. Implement corrective Actions
Check	CONTROL	Evaluate Results & Processes	7. Define Actions to aviod Recurrence
Act		Standardize Success	8. Congratulate your Team

Figure 4. Improvement processes and respective structures (Liesener, 2013)

PDCA

The PDCA cycle is an iterative qualitative tool invented by Deming for improving processes. It consists of the four phases Plan, Do, Check, and Act, starting with the Plan phase (Bergman & Klefsjö, 2010). The cycle is visualized in Figure 5.

- **Plan**: The goal of this phase is to define and clarify the problem. Begin measuring and analyzing the data as well as develop hypothesis for the problem.
- **Do**: Test the hypothesis from the plan phase and measure the results. The test is preferably performed on a smaller scale such as a pilot test to reduce cost.
- **Check**: Analyze the results from the "Do" phase and decide whether the hypothesis is true or false. If it is true one shall proceed to the next step in the cycle "Act". If the hypothesis is false, one shall return to the plan phase and formulate a new hypothesis.

- **Act**: Take action based on the result of the hypothesis and implement the solution.

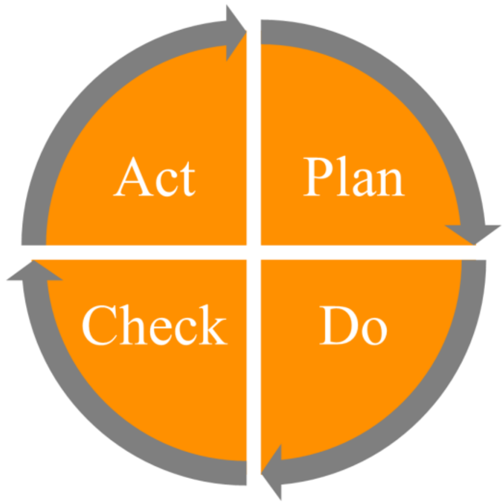

Figure 5, PDCA cycle

A3

A3 is a problem-solving methodology invented by Toyota. It utilizes an A3 sized paper as base for the report which provides a structured method for solving the problem. The method's structure can be seen below. This method is mostly used within companies working with LEAN. (Matthews, 2011)

1. The problem is first identified and clarified so everyone is in agreement of the problem
2. Break down the problem and understanding the current state.
3. Quantify the problem and start observing and measuring
4. Start conducting root cause analysis. This is supported by statistical analysis and techniques such as five whys.
5. Create counter measurements and hypothesis for solutions.
6. Implement and analyze the counter measurements developed in the previous step.
7. Evaluate the results of the counter measurements to see if the improvements have had intended effect.
8. Standardize the process if the intended effect was attained. (Matthews, 2011)

8D

8D stands for eight disciplines problem solving and is a method for structured improvement mainly used in the automotive industry (Koncz & Pokorádi , 2018). The method is composed of eight structured steps which can be seen below.

1. Assign work group and scope the problem
2. Narrow down the problem and describe it
3. Define containments action to isolate the problem
4. Find, analyze and verify the root causes with the help of statistical diagrams and five whys
5. Come up with preventive actions from the root causes found. Use FMEA to grade them

6. Implement the best solutions from the previous step
7. Measure the benefits of the solution
8. After solving the problem recognize everyone involved and congratulate each other on a job well done. (Koncz & Pokorádi , 2018)

Six Sigma and DMAIC

Six Sigma is a data driven problem solving approach to improve processes (Brook, 2017). The model was developed by Motorola in the 1980s to reduce defects in the production with the help of structured statistical analysis. The Six Sigma approach is usually visualized as a circle as the model is especially to advantage when analyzing repeating processes which can be seen in Figure 6 (Bergman & Klefsjö, 2010).

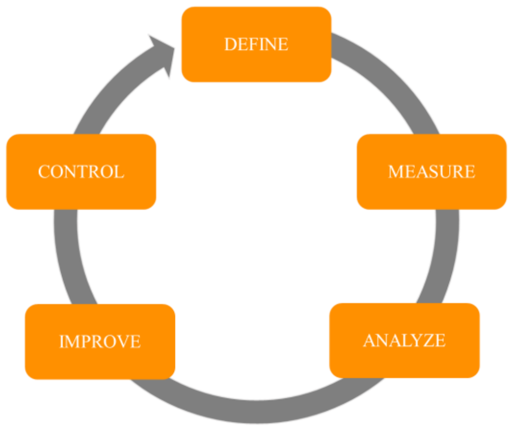

Figure 6. DMAIC cycle (Bergman & Klefsjö, 2010)

It is a stage-gate model that consists of five phases known as DMAIC, Define-Measure-Analyze-Improve-Control. Almost all projects are conducted the same way with the five succeeding phases starting with the Define phase (Barone & Franco, 2012). The aim of the phase is to define the problem, set boundaries and clarify a goal for the project. It is crucial to only focus on defining and understanding the problem rather than searching for solutions. The next phase of a Six Sigma project is the Measure phase where the goal is to narrow down the problem, collect the correct data and conduct initial measurements in the project. The third step is the Analyze phase which is conducted in three steps. The first step is to find patterns and variations in the data collected from the Measure phase. The second step is to analyze the data and find the root causes to the problems, the third and last step is to verify to ensure the root causes are correct. The fourth phase in the Six Sigma process is the Improve phase where the focus shifts from analytical to innovative. The aim is to find possible solutions for the verified root causes. The last phase in the Six Sigma process is the Control phase where the aim is to make sure that the implementations from improve have the desired effect. (Brook, 2017), (Barone & Franco, 2012), (Magnusson, et al., 2003), (George, et al., 2004). Table 2 present an overview of tools used in the phases in the DMAIC approach, for a more detailed view of tools and their use areas see (Brook, 2017).

Table 2. Six Sigma phases and tools with brief explanations (Cronemyr, 2007)

Phase	*Tool*	*Explanation*
Define	Business case	Define gap between current process performance and strategic targets; estimate financial benefit of project.
	Project charter	Set target, scope, boundaries, schedule and team members.
	SIPOC	High-level process map (Suppliers, Inputs, Process, Outputs and Customers).
	VOC	Voice of the customer. Who and what? Use the Kano model.
	CTQ driver tree	Critical to quality. Transform the voice of the customer to internal process measurements.
Measure	Measure variables	Gather all variables that could be measured.
	Prioritisation	Select most important variables to be measured.
	Data collection plan	Characterise variable type, set up measurement system, define operational definition.
	Gage R&R	Check measurement system for errors and built-in noise.
	Sampling	Perform measurement of historic and/or current data.
	Data presentation	'Eyeball data'. Usually with Pareto charts and histograms
	Process capability	Calculate initial process capability and process sigma.
Analyse	Process door analysis	With limited data, find problem areas in detailed process map analysis.
	Data door analysis	With extensive data, find problem areas in detailed data analysis.
	Cause and effect analysis	Ishikawa diagram ('fishbone diagram') to find root causes to problems.
	Regression analysis	Find transfer functions by regression of measured data.
	Hypothesis testing	Test connections/relations between input and output variables by testing of hypotheses.
	DoE	Design of Experiments. Used to get linear and nonlinear transfer functions between inputs and outputs.
Improve	Generate solutions	Innovation techniques to generate several possible solutions to root causes found in the Analysis phase.
	Cost/Benefit analysis	Remove or change solutions that are not feasible.
	Pugh matrix	Selection, evolution and synthesis of concepts (*not always used in DMAIC*).
	Selection of solution	Select the best solution that fulfils all must-be demands and high scores on more-is-better and delighter demands.
	FMEA	Failure Modes and Effects Analysis. Risk evaluation of the solution and of the implementation of the solution.
	Pilot phase	Test run of process improvement with a limited scope.
	Implementation plan	A 'normal project plan' for an implementation project. Schedule, responsibilities, *etc*.
Control	Process control chart	Set up a process control chart to be used by process owner with actions for predefined deviations.
	Documentation and standardisation	Update all documentation of the process. All old files must be removed. Train employees in new process.
	Monitoring	Monitor input and output of improved process.
	Evaluation of results	Calculate process capability and process sigma of the improved process. Compare to targets.
	Key learnings	Document key learnings from the project.
	Project closure	Publish the results of the project and celebrate with the team.

Several criticisms and limitations regarding Six Sigma exist. Common limitations raised are that initial costs for implementation of Six Sigma are too high and require substantial investment. Almost 70 % of change management initiatives fails according to estimates and Six Sigma initiatives report around a 60 % failure rate most commonly caused by lack of management commitment or lack of statistical knowledge. Due to this there is criticism that the benefits do not outweigh the efforts and cost needed for implementation. (Anthony, et al., 2019). Furthermore, there are limitations as poor implementation could negatively impact customer and employee satisfaction. However, studies suggest successful implementation promotes customer satisfaction. There are also contradicting views on how the structured and disciplined DMAIC-approach affect employee creativity and innovation, with claims that it stifles innovation whereas others claim it fosters innovation. (Anthony, et al., 2019). The knowledge of conducting root cause analysis, and in particular Six Sigma projects, are often too statistically advanced for companies or they lack the resources to conduct the projects which leads to decision making based on non-statistically proven data (Godfrey, 2008).

Summary of improvement process

The four improvement processes presented are all similar and follow a structured methodology for findings root causes and implementing solutions to them. Liesener (2013) argues that the main difference between the processes is the scale of the problem. He claims that PDCA and A3 are primarily used for medium sized problems while DMAIC focuses on large and complex problems. Further can 8D be used for many types of problems but is primarily used for more complex problems in the automotive industry (Riesenberger & Sousa , 2010). Smaller problems benefit from utilizing fast paced cost-effective methods such as Kaizen Blitz with a “just do it” approach (Liesener, 2013).

3.1.2 Statistical and Qualitative Tools for Root Cause Analysis (RCA)

Root cause analysis aims to identify the true cause of problems and the necessary actions to eliminate it, by use of a structured investigation. It is a collective term for describing different tools, approaches and techniques to detect causes of problems. Only by finding and eliminating the root cause of a problem it is possible to prevent the problem from reoccurring. (Andersen & Natland Fagerhaug, 2006). Knowledge about variation is necessary in root cause analysis to achieve lasting results, as variation is present in all processes. It is important to distinguish between natural, or random variation and variation due to identifiable causes to be able to draw conclusions about the process and finding relations between parameters. (Bergman & Klefsjö, 2010). Correlation is a statistics term for measuring relation between parameters. Xiong, et al. (2019) describes correlation as “the tendency of two variables to change together”. A correlation can range from -1 to 1, where -1 indicates a very strong negative correlation, 1 indicates a very strong positive correlation and 0 indicates no correlation (Schober, et al., 2018). However, a strong correlation does not imply causation meaning that one variable affect the other, also known as cause-effect relation (Barone & Franco, 2012).

There is a statistic relationship called spurious correlation in which two variables seem to be causally related but are not. Therefore, one must proceed with caution when analyzing bivariate data as it is easy to draw wrongful conclusion between correlation and causation (Xiong, et al., 2019). Figure 7 shows a strong correlation between violent crimes and ice cream sales, will an ice-cream ban thus reduce the number of violent crimes? Of course not. The diagram only displays two variables that correlates but does not have the capability to indicate causation between the two.

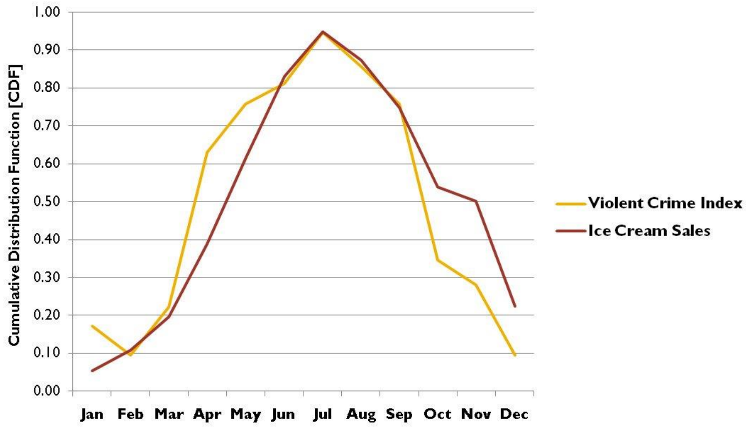

Figure 7. Graphical representation of spurious correlation (Hamra & Schinasi, 2017)

The two variables are confounded with a third part, in this case warmer weather. The heat will increase peoples' tendency to buy ice-cream as well as their desire to go outside. More people outside will in turn generate more crime. (Hamra & Schinasi, 2017). It is not the increase in ice-cream that kills, it is the warmer temperature that affects the two variables. Furthermore, spurious correlation which are not affected by a third confounding part can also be found, it is just random data which happens to correlate (Vigen, 2015).

When a relation between variables is found it is important to understand whether it is statistically significant or not. Statistical significance is a term within statistics when the difference between an observed value and a comparison value is large enough to not be by chance (Brook, 2017). It is one of the cornerstones in statistics and is central in hypothesis testing. There are several ways to test for significance depending on the data provided. For example, T-test can be used to test if there is a difference between the mean of two populations. Analysis of Variance, ANOVA, can be used to test the difference in mean or variance between three or more populations (Montgomery, 2013). However, statistical significance can emerge even though there is no causal relationship as can be seen in Figure 7. Therefore, it is important to use qualitative tools and include individuals with good process knowledge in order to understand if the variable relations are causal or not.

A commonly used qualitative tool to find root causes is "Five Whys" together with Ishikawa-diagrams. It is an iterative method where you ask the question "Why" repeatedly until a root cause is found (Brook, 2017). As the method goes in depth of each problem it is a powerful tool to find out if the problems from the statistical tests are root causes or not. A problem can have several root causes, but the same method can be used for all problems (Brook, 2017). The method has however gained criticism as the question "Why" may be misinterpreted and lead to faulty root cause assumptions.

3.2 Process Management

The main purpose of a process is to create outputs of greater value than a set of inputs through repeated activities. The output of an organization is the result of interrelated business processes (Melan, 1993). Process management is a methodology to improve processes, which stems from the importance of adopting a process view and to continuously improve (Bergman & Klefsjö, 2010). Process Management is an integrated and unified approach that produce consistent and effective results toward TQM and provides sharp focus on the manner in which a business is conducted. It forces a focus on the flow of work independent of organization. (Melan, 1993) The procedure of process management consists of four steps (Bergman & Klefsjö, 2010) presented below.

- *Organize for improvement,* where process owners and process improvement teams are appointed.
- *Understand the process,* where processes are mapped, and customers and suppliers are defined.
- *Observe the process,* where measurements are implemented, and control points established.
- *Improve the process* is the final step where feedback from measurements are analyzed to find improvements. (Bergman & Klefsjö, 2010)

3.2.1 Process Maturity

It is estimated that change management initiatives in organizations fail almost 70 % of the time (Anthony, et al., 2019). Cronemyr & Danielsson (2013) found that implementation failures often are caused by organizations not being mature enough rather than the implemented method being poor. They present a Process Management maturity model with a diagnostics tool to assess what level of maturity a process fulfills and how to move up to the next level.

The process maturity model is built up by four levels and one pre-level to describe the state where none of the maturity levels have been reached. Lower levels act as foundation for the next to describe natural steps for implementing process management. *Business need* is the pre-level and primarily symbolize a lack of maturity. No processes or process awareness within management exist and the organization does not fulfill prerequisites to establish processes. (Cronemyr & Danielsson, 2013). The model can be seen in Figure 8.

On the first level, *awareness*, an awareness of the usefulness of working with processes has been gained, mainly at management level. Awareness is necessary for commitment to establish processes. (Cronemyr & Danielsson, 2013)

At level two, *established*, the process is fully established, defined with requirements, documented, and used in practice. Attention and demanded results from active management, as well as structures to maintain and update the process are needed for sustainability of the process. (Cronemyr & Danielsson, 2013)

Level three, *improved*, is where improvements can start to take place. Only when the process is fully established can it be improved. Process data start to play a role as the improvements need to be based on facts. To achieve this, customer focused measurements and structured methods for conducting improvements are required. Controls should also be developed and established. (Cronemyr & Danielsson, 2013)

The fourth and final level, *adapted*, requires commitment and understanding by everyone using the process. The process works in collaboration with the customer, is more flexible and proactive and can adapt to specific customer requirements. It is possible to react to changes sooner and tune the process by use of statistical process control. (Cronemyr & Danielsson, 2013)

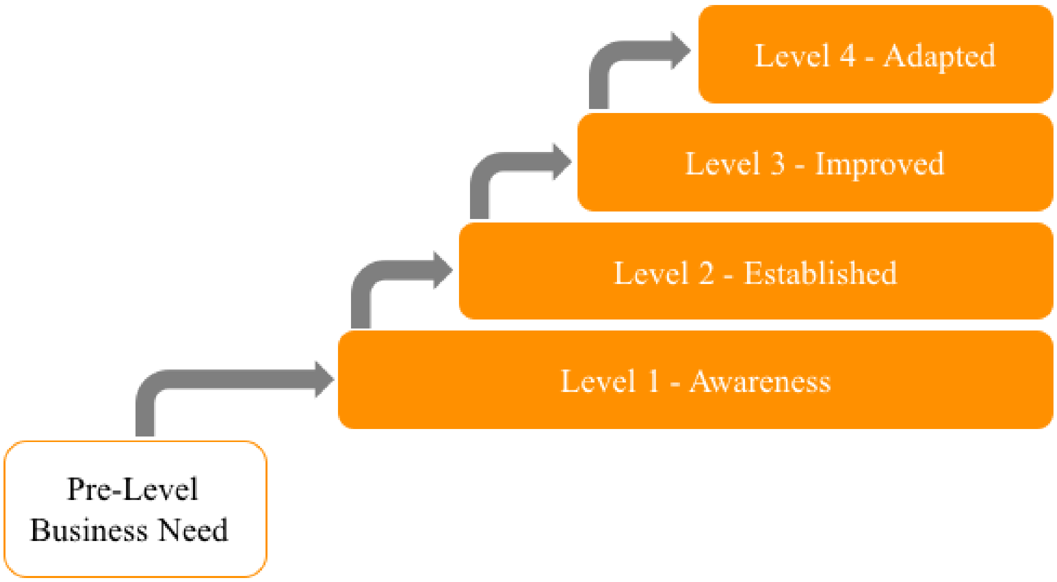

Figure 8 Process maturity model (Cronemyr & Danielsson, 2013)

To determine what level a process is currently at, two questionnaires, one for process owners and one for process users, are used. To avoid biased interpretation of answers, there are pre-determined answer alternatives to questions divided according to eight different categories, presented in the list to the left in Figure 9. For the blank spaces in Figure 9, the same criteria as the previous level is used. Respondent's results are compiled separately and then combined in an overview category-wise. The category that is considered the weakest link determines the overall process maturity. Future improvement initiatives can therefore be prioritized because of the categorical overview provided. (Cronemyr & Danielsson, 2013)

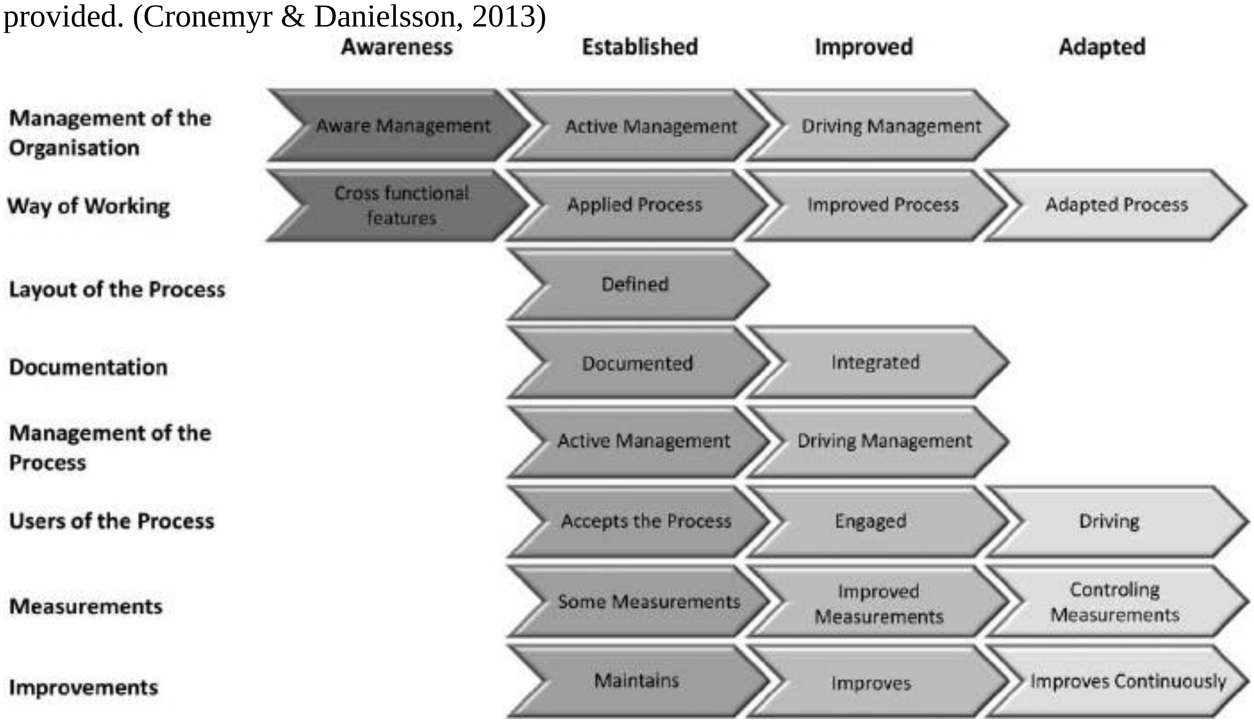

Figure 9. Diagnostic tool, cited in (Cronemyr & Danielsson, 2013), based on (Bergholtz & Danielsson, 2012)

3.2.2 Digitalization and Process Management

Sehlin, et al., (2019) developed a conceptual cooperative model for small and medium enterprises to digitalize their business processes in response to companies failing to connect their core processes to new digital business processes. The model incorporates processes, digitalization and innovation and start with establishing a *current state*, where it is evaluated how an organization currently work within the three areas. *Prioritization* is then performed as step two, based on key performance indicators, (KPIs)and process maturity, to get a better understanding which processes to start digitalizing. In step three, *digital roadmap,* the purpose is to find optimal digital solutions to a prioritized process through a structured path. The last step involves establishment and implementation of selected digital solutions. The authors found that there is a connection between process maturity and digital solutions, and it indicated that processes with higher process maturity can meet digital opportunities more efficient. The authors also conclude that the process maturity level *awareness* imply digital solutions related to digitizing and core offerings. (Sehlin, et al., 2019)

3.3 Digital Transformation and Digital Maturity

Digital transformation has many definitions, in a literature review Teichert (2019) highlights two different ones often used. The first being *a disruptive or incremental change process that starts with adoption and use of digital technologies and evolving into an implicit holistic transformation of an organization*. The second definition used is *an evolutionary process leveraging digital technologies to enable business models, operational processes and customer experiences to create value* (Teichert, 2019). The best understanding of digital transformation according to Kane (2017), is adopting business processes and practices that help to compete effectively in an increasingly digital world. This means that the implementation of new technology is only one part of the digital transformation and it is not hard to find companies that have implemented a new digital tool just to have it unused by employees or unable to deliver a transformational impact (Kane, 2017). He also argues that managers should shift focus from digital transformation to maturity as there are always ways to become more digitally mature and the process is never complete (Kane, 2017). Chanias & Hess (2016) define digital maturity in two different ways, technological and managerial. Technologically it could be described as to what extent a company's tasks are performed and information are handled by IT. The managerial view is the status of the digital transformation of a company, describing what has been achieved in performing transformation efforts. (Chanias & Hess, 2016)

3.3.1 Industry 4.0 and Internet of Things (IoT)

Industry 4.0 is a complex term which has no real definition but has its base in digitalization and automation. The term aims to explain the increasing connectivity and measurability with IoT devices, sensors and big data where everything is interconnected (Alp & Emre, 2018). IoTs describe a network of devices connected through the internet, that are able to collect and exchange data using embedded sensors. The digital world is becoming more quantifiable and accessible through the data obtained by the proliferation of connected sensors. (Alp & Emre, 2018)

3.3.2 Analytics

Phil Simon (2017) defines analytics as the process of going from raw data to valuable insights and increased understanding of a topic. Sharda et al.(2014) describes it as the process of developing actionable decisions or recommendations for actions based on insights from historical data. Simon (2017) also states that ideally, analytics explains why something happened and the best analytics suggest corrective and measurable actions. Analytics are generally categorized as either *descriptive, predictive, or prescriptive*. Descriptive analytics summarize or describe what is happening or has happened (Simon, 2017). Predictive analytics attempts to forecast what will happen based on past performance and detect patterns and relationships in historical data (Evans, 2012). Techniques that fall under predictive analytics are statistical or data mining (Sharda, et al., 2014). The goal of prescriptive analytics is to provide decisions or recommendations to optimize the performance of a system (Sharda, et al., 2014). Sometimes *diagnostic* analytics is used to describe analytics answering the question why something happened, with techniques such as drill-down, data discovery and correlations. (Gartner, Inc. , 2020)

3.3.3 Big Data Analytics

The term big data itself is a relatively new concept which arose with the potential to store and process information digitally. However, the term is vaguely defined and has become a ubiquitous lingo for anything associated with large datasets. (Boyd & Crawford, 2011). De Mauro, et al. (2014) attempts to establish a conclusive definition of the term by analyzing articles within the subject. He identifies three key characteristics when defining the term big data, these are "Volume, "Velocity" and "Variety" (De Mauro, et al., 2014). Volume signifies the quantity of the dataset. Velocity stands for the speed in which the dataset is changing. Lastly Variety refers to the diversity in the data. This includes everything from the format to the structure of the data. (EMC education Services, 2015).

Big Data Analytics (BDA) is a concept used for advanced analytics of enormous data sets. The use of BDA can help to find patterns and analyze data which is impossible for humans in order to assist in decision making. There are however backsides of Big Data Analytics. Calude and Longo (2016) argues that when handling colossal amounts of data there is always the risk of introducing spurious correlation due to the sheer amount of data. To perform successful big data analysis, Su (2019) argues that understanding the engineering problem is crucial and present three key principles to consider. These principles are *data quality*, relating to missing values, incorrect records or problems related to measurement. *Methods of data analysis*, how to effectively process massive amounts of data and use appropriate tools to extract information. The last principle is *customer perspective*, whether or not a problem is a customer concern. (Su, 2019)

3.3.4 Data Mining

Data mining is the study of collecting, cleaning, processing, and analyzing data to gain useful insights. It is a complex process with several stages such as data collection, preprocessing and analysis. Real encountered applications vary in terms of problem domains, formulations, and data representations. It is difficult to create reusable and general techniques as each data mining application is unique. The workflow of a typical data mining process is divided into three stages, starting with data collection. To collect data specialized hardware, manual labor or software tools are often required, this is then stored in a data warehouse for processing. Feature extraction and data cleaning is the second step, where data collected often need to be often integrated and transformed from multiple sources into a unified format. The final part is analytical processing and algorithms, where analytical methods from processed data is designed. (Aggarwal, 2015)

3.3.5 Data Warehouse

Data Warehouse is a system that gathers data from dispersed data sources within an organization. The idea behind it is to enable analysis and reports between different data sets. A data warehouse is a pre-requisite for more complex analysis involving big data and is often used as a backbone for Business Intelligence (Vaisman & Zimányi, 2014). By combining data sets hidden correlations may be detected and more powerful analysis can be performed (Jukić & Nestorov, 2006).

3.3.6 Data Cleaning

Data cleaning, also known as data cleansing, is one of the biggest challenges in data analytics (Chu , et al., 2016). It is an umbrella term for detecting, removing, and structuring datasets and to ensure a higher data quality (Van den Broeck, et al., 2006). It is primarily used in lower quality datasets but have with the rise of digitalization and big data seen an increase in all kinds of datasets (Tang, 2014). Tang (2014) argues that up to 30% of a company's data is unclean which can be extremely costly as wrongful decisions are based on this data. No matter how well the data is collected there will always be room for error which will need to be cleaned. Errors include duplicates, misspellings, corrupted data, or mixed formats. The cleaner the data the more accurate the analysis will be. However, it is a debated subject as some claim that it is data manipulation meanwhile other see it as a necessary step in order to correctly analyze the data (Van den Broeck, et al., 2006).

3.3.7 CRISP-DM

CRISP-DM is an acronym for *Cross-industry standard process for data mining* and it is the most widely used framework for data mining projects. It is a model describing common approaches used in data mining projects. The framework consists of six phases often used iteratively.

The first phase, *business understanding*, is about translating objectives and requirements into a data problem definition and preliminary plan. In the second phase, *data understanding*, is data collected and activities to familiarize with the data to identify quality problems and get first insights into the data performed. *Data preparation* are the activities necessary for constructing a final data set used by modeling tools. During the *modeling* phase parameters are calibrated towards optimal values. When moving in to the fifth phase, *evaluation*, a model that seem to have high quality for data analysis have been built. The model is evaluated based on criteria such as steps executed to create it, or if some business perspective has not been addressed sufficiently. *Deployment* is the last phase where the project is handed over to the client. (Chapman, et al., 2000)

3.3.8 Data Quality

Data quality is often defined by the six quality dimensions, *Completeness*, *Consistency*, *Uniqueness*, *Validity, Timeliness* and *Accuracy* (DAMA UK Working Group, 2013).

Completeness represents to what degree the required data is recorded and available in our database compared to the real world. Required data in this case is what data that is needed to perform an analysis. Completeness dimension can be very hard to fix after the data is collected but can be done by asking the respondents to fill in the questionnaire again or re-interviewing depending on the data collection technique (Dasu & Johnson, 2003), (DAMA UK Working Group, 2013).

Uniqueness measures the ability to only record each unique data point one time. For example, Two different databases including European capitals are analyzed and both *Lisbon* and *Lisboa* are included as unique values even though they are the same city (Dasu & Johnson, 2003), (DAMA UK Working Group, 2013).

Timeliness represents to how recent the data was collected or updated and used. There will always be a gap between the collection and the use of data, therefore it is important to observe the frequency of change in the data. Some data needs to be analyzed more carefully as they may change over time while other are constant over time (Dasu & Johnson, 2003), (DAMA UK Working Group, 2013).

Validity refers to the conformity of the set rules and constraints for the data and its collection. For example, constraints regarding the datatype, the range of the data or certain patterns such as telephone numbers starting with a plus sign and being separated by a hyphen after three digits (Dasu & Johnson, 2003), (DAMA UK Working Group, 2013).

Accuracy measures how far the value in the database is from the true value. This is also hard to measure as the true value is not always known. For example, a person claims to be a certain weight which is impossible to examine as there is no true value to double check against (Dasu & Johnson, 2003), (DAMA UK Working Group, 2013).

Consistency refers to the similarity within and between datasets. For example, a customer might have two different genders in different databases whereas only one is correct. It can be difficult to correct consistency but there are methods such as choosing the latest recorded, the most trustworthy or simply asking the customer (Dasu & Johnson, 2003), (DAMA UK Working Group, 2013).

3.4 Digital Tools and Techniques

Digital tools are a broad term which is defined differently depending on who you ask. Steils and Hanine (2019) describes the term as "Tools characterized by electronic and especially computerized technologies" which will also be the definition used in this thesis. Examples of digital tools includes everything from mobile phones, IoT sensors, statistical software, and programming techniques such as data mining and machine learning.

3.4.1 Drawing and Modelling Tools

Drawing and modelling tools are computerized illustration programs which is used for visualization. One of its strength is visualizing processes and displaying connections between them. The tools can visualize levels in the processes, such as main and support processes and how they are related. Modelling tools are commonly used by organizations to digitize and visualize processes and their respective activities. (Propia, 2020)

3.4.2 Spreadsheets

Spreadsheets are amongst the most popular and widely used software and is a type of computer software which is used to perform mathematical calculations. The program is composed by cells in which data can be entered. Pre-defined calculations can then be performed with the data in the different cells. The most known Spreadsheet program is Microsoft Excel. (Broman & Woo, 2018)

3.4.3 Statistical Software

A statistical software is a computer software which is niched towards more advanced statistical computations. Common calculations include normality test, analysis of variance, regression models and more. The software is usually used when performing analysis on larger datasets which require in depth examination. There is a vast amount of statistical software on the market, examples are MINITAB, SPSS and SAS. (Ali & Bhaskar, 2016)

3.4.4 Electronic Brainstorming

Brainstorming is a well-known technique which is used to generate as many solutions as possible to a problem. There are some basic rules which include, no criticism of ideas, all ideas are acceptable and use other ideas to come up with more. There are however some drawbacks with traditional brainstorming which electronic brainstorming solves, the first being blocking. This occurs when you have to wait to present your idea because someone else is presenting which may result in you forgetting yours. With electronic brainstorming, everyone can submit their ideas at any time. Electronic brainstorming solves this by letting the participants post their ideas simultaneously. The second shortcoming is evaluation apprehension. It has been shown that participants in traditional brainstorming actively withhold ideas due to being anxious of others evaluation of your ideas. This is nullified by Electronic Brainstorming as all proposals are anonymous. (Gallupe & Cooper, 1993)

3.4.5 Business Intelligence Platform (BIP)

Business Intelligence (BI) has developed into a buzzword and a new trend within business improvement. The term was however introduced in the 1990s as an umbrella term for methods and techniques used for collecting and analyzing data. Over time the expression has changed and is today often described as a data driven decision support system. The term includes data gathering, databases management, visualization tools and analytical tools (Gray & Negash, 2008) (Azeroual & Theel, 2019). Recently an addition of the words “System” or “Platform” has been added to describe the newer decision support systems to differentiate from the original umbrella term.

A Business Intelligence Platform is a software used for displaying and visualizing huge amounts of data gathered from different sources which people are incapable to do (Sharda, et al., 2014). They are doing it by enabling an interactive environment where the user can drag and drop pre-programmed charts and analysis tools such as regression models which requires no previous programming knowledge or statistical background (Gray & Negash, 2008). That is however also the backside of BI platforms, the tools can tell people what is happening but not why it is happening (Gray & Negash, 2008). The user interpreting the results needs to understand the data and statistics behind it to draw the correct conclusions (Eckerson, 2007). Duarte (2017) argues that BI software packages provide a localized view of many variables but is lacking when it comes to providing an end-to-end view of a process. BI platforms are especially useful for *drill down*, which is enabled by starting from a high level KPI and then follow a data trail through different levels of detail. There are several popular BIP on the market which serves the same purpose, the most known are Microsoft Power BI and Qlik Sense. (Azeroual & Theel, 2019)

3.4.6 Process Mining

The term Process Mining is a technique for mapping and analyzing processes, traditionally done by using workshops. This will however result in an idealization instead of a realistic depiction of the process. With modern technology, this can instead be done digitally with the use of process mining. The technique utilizes real data from the process event logs in order to map out and analyze the process. (Alvarado, et al., 2012)

To use process mining an event log where each row correspond to an event is required, it also must consist of at least three columns, containing a case id, timestamp, and activity name. Process mining is usually divided into three categories, *process discovery*, *conformance checking* and *enhancement*. A discovery technique extracts data from an event log and produce a process model. Conformance checking compares an existing process model with an event log of the same process. With enhancement, the idea is to improve an existing process model using the information recorded in an event log. (van der Aalst, 2016)

Claes & Poels (2012) highlights the benefits and drawbacks of process mining by conducting a survey. They found that biggest assets were a mix of characteristics and applications. The main advantage presented was objectivity in the data followed by conformance checking and the possibility to find root causes to problems. Furthermore, the biggest drawbacks were mostly characteristics related to data access, cost of implementing and lack of data quality. There were also some respondents who raised concern regarding the lack of guidance and difficulties in understanding the software.

3.4.7 Text Mining

Text mining is a technique closely related to data mining which is utilized to extract relevant information from written sources. This includes everything from emails, books and websites to reviews etc. Companies can by exploiting text mining access information which otherwise would be undetected. This will lead to better understanding and more accurate data-driven decisions. The underlying method is based on several statistic methods such as clustering and categorization. (Sun & Cai, 2009)

3.4.8 Digital Twin

Digital Twin is a concept of creating a digital copy of a process, product or service. This enables analysis and testing of real time data in order to generate preemptive action towards problems before they occur. The technology has just recently gained attention due to the need of massive amount of real time data required to use the technology. The upsurge of Internet of Things (IoT) has however increased the use of digital twins as it produces sufficient amounts of data. (Kritzinger, et al., 2018)

3.4.9 Discrete Event Simulation

Discrete Event Simulation (DES) is a method of mimicking and testing real processes and systems in a computerized setting. A DES software makes it possible to add “what if” scenarios as well as change statistical variations in real processes that transpires over time. The tool enables an inexpensive way of testing changes to a process and how it affects other process or the system as a whole. (Mansharamani, 1997)

3.4.10 Machine Learning

Machine learning is a technique and considered a subgroup to Artificial Intelligence. The idea behind it is that machines can self-learn from data with little to no intervention from humans. The technique is based on finding patterns in the data and drawing conclusions based on past events to predict the future. The technique is for example used when giving recommendations on Netflix based on earlier preferences or diagnosing bone fractures on x-rays or perfecting automated self-driving cars. (Selvam & Babu, 2015)

3.4.11 Robotic Process Automation

Robotic Process Automation, RPA, is an umbrella term for tools operating on the user interface of computer systems in the same manner a human does. RPA tools aim to reduce the burden of simple and repetitive tasks by mapping a process for a software robot to follow and performing - if, then and else- statements on structured data. RPA does not change the existing systems but rather operates on top of them. (van der Aalst, et al., 2018).

RPA application is primarily beneficial to standardized processes and the first step for a process in RPA is therefore a harmonization and standardization of processes to understand the maturity of a process. Trying to use RPA on a process with many variants rarely generate beneficial results as the investment outweighs the savings. Process mining can therefore greatly benefit RPA by visualizing the actual process flow and the variations that occur. (Geyer-Klingeberg, et al., 2018)

3.5 Integration of Digital Tools in Improvement Processes

Common characteristics in why and how digital technology has been applied in synergy with structured improvement work are presented below. Several of the studies utilize a methodology based on the DMAIC structure.

Motivation for approach with integration of digital tools in improvement processes

Modern datasets have grown in size and complexity and because of this, traditional statistical methods (e.g. t-test and linear regression) are less effective (Zwetsloot, et al., 2018). Such methods are also incapable of modeling the effect of both categorical and continuous predictors and depend on assumptions of normality (Ghosh & Maiti, 2014). Fahey et al. (2020) motivate their use of an integrated approach because normality assumptions do not hold for the modern biopharmaceutical manufacturing dataset and its size and complexity makes it difficult to analyze with traditional methods. Analytics tools can also generate hypotheses from data whereas Six Sigma use data to confirm hypotheses by process experts (Fahey, et al., 2020).

Ghosh & Maiti (2014) argues that current tools for root cause analysis in Six Sigma, such as Ishikawa and Pareto charts, cannot identify complex relationships or interactions among impacting variables. Design of Experiments (DOE), often used to confirm root causes, also has a restricted application because of the possibility to only test a few parameters at the same time and production need to be disrupted to perform experiments (Ghosh & Maiti, 2014). Schäfer et al. (2018) support this view by highlighting the capacity to test more parameters with the use of data mining tools instead of a classical experimental design with few parameters.

Schäfer et al (2018) see potential for process improvements, a key quality management area, by the use of data mining as the data proliferation are increasing. Fahmy et al., (2017) start from a specific business problem with the purpose of showing how a data mining approach can enhance Six Sigma problem solving. For comparison of results, a regular Six Sigma project is run in parallel with the project including data mining and machine learning techniques.

Zwetsloot et al. (2018) investigates three separate cases where the motivation for integration of data science has been that traditional improvement projects have failed. Gupta et al (2020) performed a systematic literature review to provide further research directions as the increase in data causes a need for BDA in Lean Six Sigma (LSS), to improve decision making and more accurate processes. A framework of potential applications of big data techniques in each phase of LSS is presented (Gupta, et al., 2020).

Applications of digital tools with quality management tools

Alvarado, et al. (2012) use the process mining step discovery to evaluate current process performance by extracting a process model based on an event log. Descriptive statistics, such as box plots, from the same event log are then used to analyze key process indicators. Data mining analysis through a decision tree are then used on issues identified in the descriptive statistics for a more detailed diagnosis on a sub process. To carry out further root cause analysis an Ishikawa diagram is constructed to support the data mining analysis. A simulation model is built based on the model from process discovery and simulated to look at current performance. Generated improvement ideas from the RCA are put into the simulation model and then simulated in different scenarios to establish a what-if analysis. The scenario with the best simulated results is then implemented. (Alvarado, et al., 2012).

Ghosh & Maiti (2014) use data mining together with quality management visualization tools in the Measure phase to interpret process behavior and narrowing the scope of the analysis. Before the data mining analysis to find significant variables is started, Ishikawa diagrams and fault trees are used to brainstorm parameters that could have impact on the defects. In the Analyze phase decision trees are constructed to find the most significant variables as well as their optimal setting to use for improvements. (Ghosh & Maiti, 2014).

Fahey, et al. (2020) makes use of data mining to construct a Variable Importance Plot which is considered an upgrade to a pareto chart. It visualizes variables in order of importance on a target and also to identify each process step's contributed variation. To generate possible root causes Ishikawa diagrams are used, and some are analyzed with e.g. hypothesis testing. If the model created describe the process behavior accurately, root causes that would normally require physical experiments to analyze, can instead be simulated into the model to get a faster and cheaper approach. In the case study, root causes for variability residing outside the dataset are also found when the model diverged from actual values. A faster pace of root cause identification compared to small-scale experiments are highlighted as a significant benefit. (Fahey, et al., 2020)

All cases presented by Zwetsloot, et al. (2018) use data mining in an iterative analyze phase as a way to segment customers into clusters for more customer targeted actions. One case utilizes data mining to narrow down the scope of the project and then further for clustering customers. In another case a classification model was built for segmentation to make predictions on which customers might fall behind on payments. (Zwetsloot, et al., 2018)

Fahmy et al. (2017) use data mining throughout the project, starting with business and data mining problem definitions in Define. In the Analyze phase prioritization techniques are used to find major vital problems, and hypotheses are tested through application of statistical and data mining tools. In the Improve phase a predictive model is used in order to be able to take preventive actions for possible defects. (Fahmy, et al., 2017).

Schäfer, et al.(2018) use QM tools to support the data mining activities such as a high-level process map Supplier-Input-Process-Output-Customer (SIPOC), to faster obtain required data and Ishikawa diagram to provide insight on contributing factors. QM visualization tools are used to gain insight into the data. A prediction model is built with machine learning techniques and applied as an error forecasting system in electronics production. (Schäfer, et al., 2018)

Duarte (2017) discuss several disruptive technologies but highlights three in particular to replace or enhance existing quality tools. These are machine learning, to replace Design of Experiments (DOE), text mining to enhance the voice of the customer (VOC) and discrete event simulation (DES) to assist in analysis of value stream mapping (VSM) and theory of constraints to identify bottlenecks. He also highlights that the use of BI software packages provides user friendly data aggregation and viewing, but only as analytical consideration.

Challenges with new approaches

Root causes hiding outside dataset are still found through qualitative tools in a cause and effect session. A barrier to implement new solutions discussed is that professionals overseeing complex processes tends toward caution when implementing changes due to potential of unforeseen consequences. (Fahey, et al., 2020). Data preparation activities and modeling connected to the Analyze phase is far more complex than regular Six Sigma because of the different techniques and algorithms used (Schäfer, et al., 2018). To create an analysis and deploying it, and putting information to use is complex, when talking about real time monitoring and fine tuning a predictive model. Building predictive models from a data warehouse is an analytical luxury. (Duarte, 2017)

Six Sigma Black Belts need to translate the business problems into data science questions to solve, because of this, an emphasis is put on team updates for improvement projects and extensions in Black Belt training to provide an overview knowledge into data science topics. Data scientists are necessary to include in improvement projects because of needed skills such as dealing with large volumes of data, data wrangling and applying various algorithms and iteratively adjusting and improving data models. It is also highlighted that the proposed project model should be used only on a certain type of problem, high complexity problems that are data rich. Further, in Figure 10, differentiators to distinguish between different projects is plotted. (Zwetsloot, et al., 2018).

SOLUTION		
Known	Unknown	

DATA VOLUME		
No Data Necessary	Low	Large

COMPLEXITY			
Low	**Just-do-it** Who will address it? By when?	**Problem Solving** Why did it happen?	**Data Analysis** What does the data suggest to do?
High	**Lean (Kaizen) Event** How should we implement the solution?	**Lean Six Sigma** What is the solution?	**Data Science in Lean Six Sigma** What is the solution using the data available?

Figure 10. Project selection matrix (Zwetsloot, et al., 2018)

4 Theoretical Analysis

This chapter is based on the frame of reference with author's interpretation which compose the basis for a theoretically based framework presented in Figure 11. This framework is then further analyzed from a perspective that includes challenges of using the digital tools and technologies for different organizations.

4.1 Initial Framework

Based on the frame of reference an initial theoretical framework was developed, presented in Figure 11. The framework showcases possibilities and gaps of integrating digital tools into the Six Sigma methodology. It presents which digital tools that can supplement or replace traditional tools and try to indicate what competences that are needed.

Several studies highlight the importance of an integrated CRISP-DM and Six Sigma methodology to get a more detailed method for data mining, as one of the main flaws with the CRISP-DM model is the lack of structure. The structured approach of DMAIC and Six Sigma quality tools can support projects focusing on data mining. Therefore, the CRISP-DM model was integrated in the Six Sigma methodology and the framework displays how the two models could work in parallel.

The framework is intended to be used as a guideline to identify feasible digital tools that can be used as supplements or substitutes of existing techniques It is important to understand the company's business needs as the framework aims to highlight possible improvements of existing tools. The framework also serves to showcase the complexity and many different application opportunities that digital tools provide. A breakdown by the DMAIC phases is then presented with further explanations of specific use cases from literature.

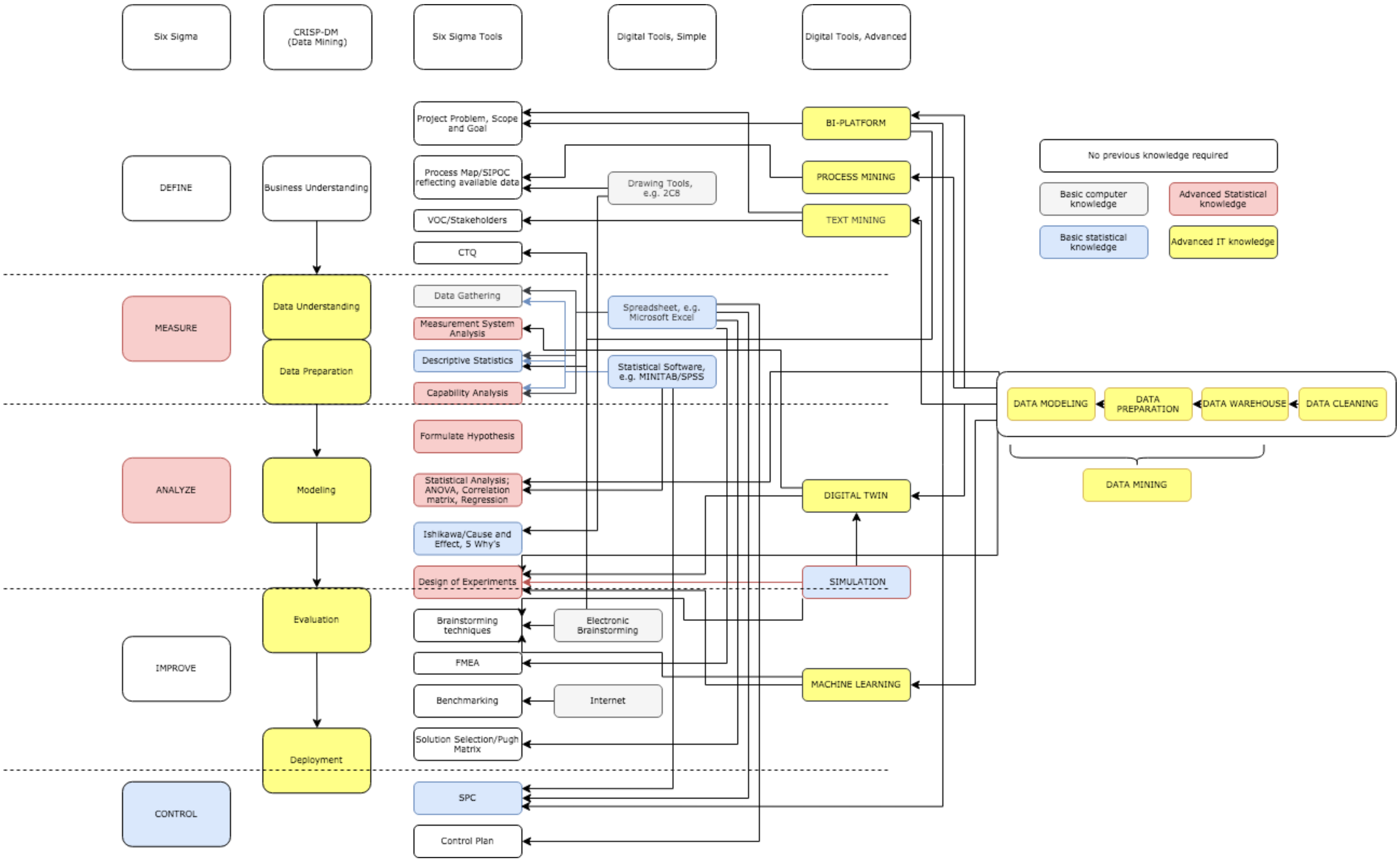

Figure 11. Theoretical framework based on literature study, author's interpretation

4.1.1 Phase by phase breakdown of framework

In the chapter below supporting explanations for the framework are presented to clarify how digital tools work as a substitute or complement to the traditional tools. In this chapter the phases of CRISP-DM will be neglected as the model is primarily integrated in the large framework to showcase where Six Sigma tools could enhance data mining projects.

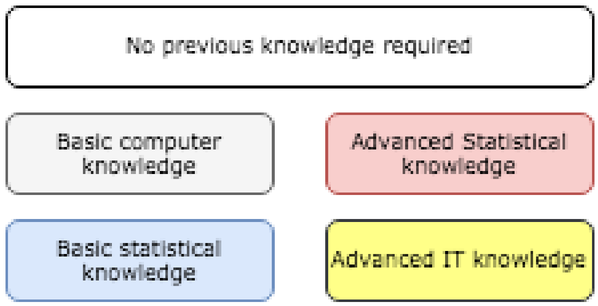

Figure 12, Color coding in the framework

Define Phase

In the Define phase the problem is clarified in several terms, how it relates to the customer, why it is a problem and what process the problem is related to. The problem statement should also be translated from a business problem into a data or process mining problem, dependent on what is applicable for the study. It is important to reflect on what available data there already is and what the project is aimed to achieve to assess which techniques are applicable. The dashed arrows in Figure 13 symbolize that the tools often collect data from a data warehouse.

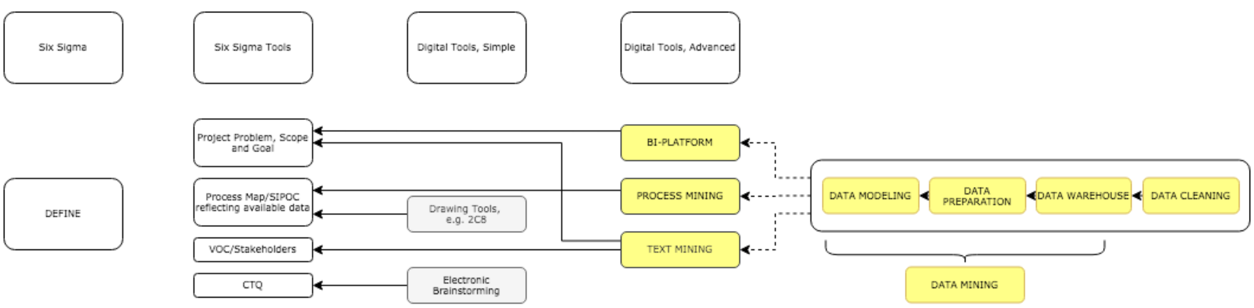

Figure 13, Define phase with digital tools

Drawing or modelling tools can be used without any prior knowledge and can assist in digitizing process maps or a SIPOC. Further can the SIPOC help to pinpoint where to look for the relevant data by highlighting relations between departments and people in the process. The discovery step of process mining can automatically create a process model from an event log to visualize the process as it happens in the system. BI software can be a useful tool in the Define phase for finding anomalies to initiate projects or help define the scope of a project. Text mining can help to get a more accurate Voice of the Customer (VoC) or to get a better problem statement.

Measure Phase

The Measure phase aims to understand process behavior and current performance. Rather than initiating data collection procedures, pinpointing which systems data need to be gathered from is a crucial activity as often data is available which save time compared to initiating a sampling procedure. Data needed are very different depending on the goal of the project and what application it is for. Figure 14 shows an overview of the Measure phase with its activities and possible tools to utilize with dashed lines to symbolize the common use of a data warehouse.

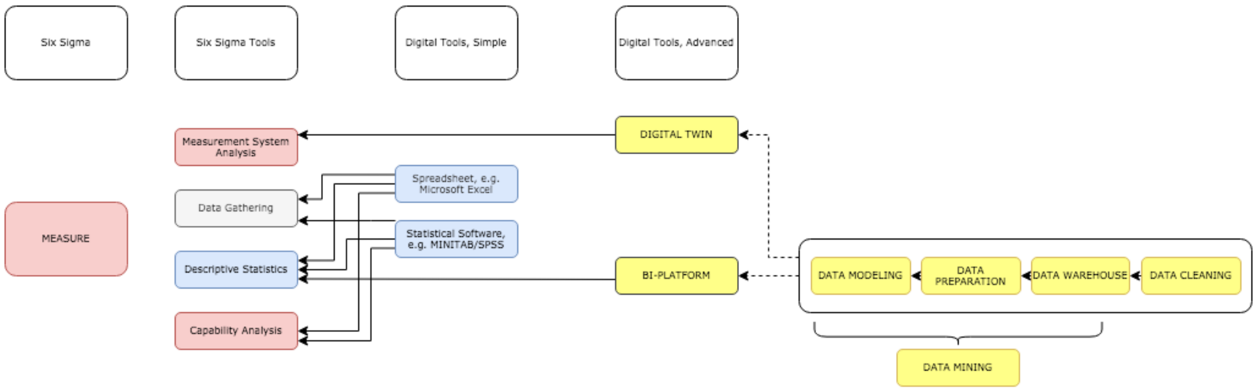

Figure 14, Measure phase with digital tools

Spreadsheets are often used to format smaller data sets and can be used for visualizing data and calculations. However, statistical software is more widely used in Six Sigma projects for this purpose as it simplifies the constructing of graphs. BI software can help to narrow down the problem by visualization and drill down in the data. Descriptive statistics provide insight to the data to narrow down where to use data mining for a further drill down in the data. In projects with more extensive data mining in the Analyze phase, Ishikawa diagrams as well as a discovered process model can be used to narrow down potential process parameters to investigate. Further can conformance checking can be used to measure the performance of a process by comparing a process model with an event log and visualize deviations that occur. Data mining can provide descriptive statistics from several integrated datasets and produce enhanced visualizations such as a variable importance plot to show the impact of different parameters on a target variable.

Analyze Phase

The Analyze phase aims to identify and confirm cause and effect relations, problems, and the respective root causes. Potential root causes are often found through Ishikawa diagrams and 5 Whys and confirmed with hypothesis testing, ANOVA or DOE. The interaction between activities and digital tools can be seen in Figure 15.

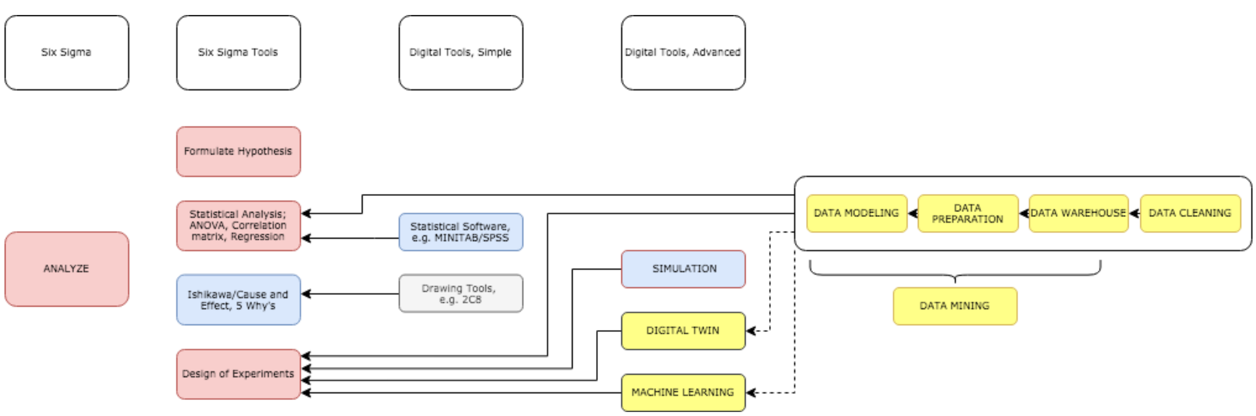

Figure 15, Analyze phase with digital tools

Depending on the goal of the project, different data mining techniques such as decision trees, clustering, association, and classification can provide segmentations or show unintuitive correlations to investigate for causality. Critical factors impacting a target variable can also be identified through data mining and either reduce parameters or eliminate the need to run a designed experiment. Instead of running a physical experiment, simulation can be a powerful tool to test hypotheses regarding process behavior or to analyze a VSM future state. If a digital twin is set up, credible simulations can be run by altering parameters in that model.

Improve Phase

In the Improve phase solutions are developed, often through brainstorming, then compared and assessed to find and implement the best ones. Assessment of solutions with FMEA can be digitized through spreadsheets and calculations simplified. Electronic brainstorming can complement workshops to increase efficiency and make sure everyone's ideas gets put forward.

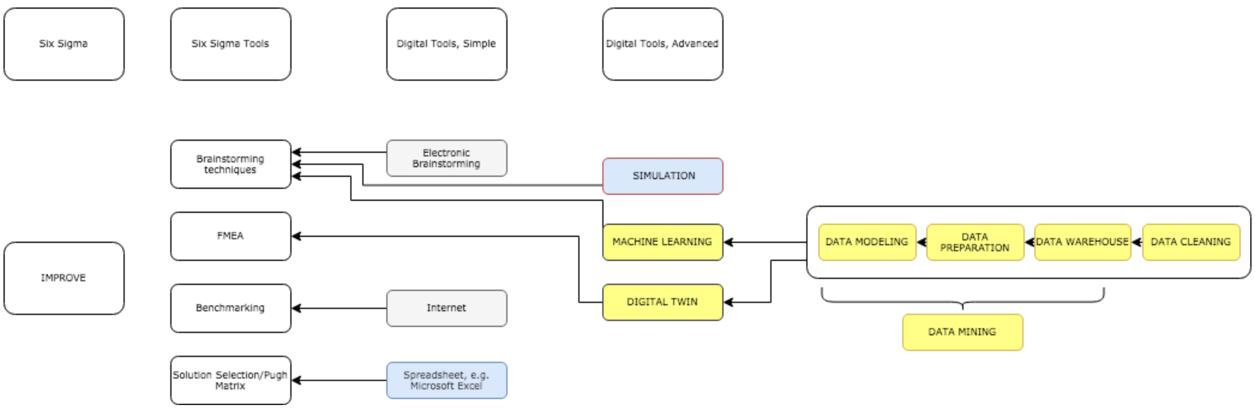

Figure 16, Improve phase with digital tools

Through some of the data mining activities (e.g. decision trees), optimal parameter settings can be found and eliminate brainstorming needs and be used for targeted process improvements. Simulation results can provide information on where to locate improvements by simulating what if scenarios with improvement suggestions implemented in the simulation model. Predictive algorithms proven to be accurate enough and can forecast defects can indicate where to put in preventive actions which can be seen in Figure 16.

Control Phase

The purpose of the control phase is to ensure that the implemented solutions become a natural part of the process. This is done by implementing continuing measurements and standardizing the solutions into the process with a control plan. Statistical process control is set up to monitor the process, most often with the use of statistical software or a spreadsheet program.

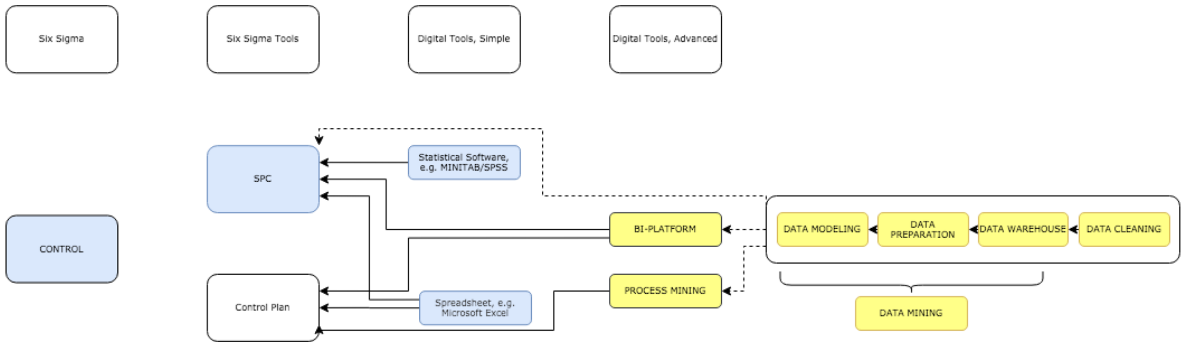

Figure 17, Control phase with digital tools

BI software can help to track the process as it goes on, with control charts embedded into an automatically updated dashboard. Predictive algorithms set up during the project can continue to run as new data is produced to monitor the process performance. Process mining can help with conformance checking which can be run continuously to assess if the process deviates from the process model. These measures can then help to initiate new projects. The likes of BI-platforms, process mining, and SPC can all be helped by a data warehouse which automatically feeds data to them, this is however not a requirement as they can function without one which is showcased by the dotted lines in Figure 17.

Summary of initial framework

Several of the digital tools has many different application areas, some of the tools are also often implemented company-wide rather using them in separate projects. BI platforms and Process Mining software are examples of such tools as the application areas are somewhat similar for most processes. Specifically, BI is often integrated into an entire organization as it provides easier access to data for process users and to be used for small analyses in the day to day operations. An important initial step in improvement projects is therefore to assess what is possible to analyze with the available data and software.

Simulation efforts can be greatly benefitted from the more accessible data, in particular discovered models from process mining as this significantly reduces time and effort needed to construct a satisfactory accurate simulation model. Data mining and machine learning applications are instead much more dependent on the specific case and goal. A model is built and trained on data sets specifically connected to the concerned process and various algorithms might have different accuracy on different processes. These applications therefore require more knowledge in data handling and programming than software intended to be integrated into an organization. Several software providers are starting to implement data mining drag-and-drop functions into its interfaces. However, knowledge of the various algorithms’ application areas is still needed.

4.2 Challenges in using the framework

The framework in Figure 11 does not display the threshold needed to overcome to implement or use the tools nor how big of a benefit it would yield compared to traditional tools. Therefore, a matrix was constructed that presents the digital tools individually and ranks them according to how much value it adds compared to the traditional tool as well as how easy they are to implement.

The *added value* column grades how much value the digital tools would add to the specific task. The rank goes from one to four where one corresponds to no value-added except for digitizing, and four relates to a high value addition. The *threshold* instead ranks on a scale A to D, how difficult it would be to implement the digital tool in the specific tasks. An A means that the tools are easy to use and requires no prior knowledge, whereas rank D requires a lot of time due to a high threshold. Moreover, a color code represents if any prior knowledge is needed in order to implement it the different digital tools.

The *added value* and *threshold* grades as well as knowledge color code can be seen in Figure 18. The grades and color codes are based on the literature and the authors experience with the tools. The rankings of the digital tools according to *Added Value, Threshold*, and *Required Knowledge* can be seen in the matrix in Figure 19.

Rank	Added value	Points
1	Does not add any value except for digitizing	1
2	Adds value, assists manual work task	2
3	Improves work tasks a lot, mix of automatic and manual work	3
4	Automatic work, requires no manual assistance	4

Rank	Threshold	Points
A	Easy to use, requires no prior knowledge	4
B	Small threshold, easy to use afterwards	3
C	Moderate threshold, requires time for effective use. Possible to self-teach	2
D	High threshold, requires a lot of time to learn. Consider buying the competence	1

Required knowledge
Basic Computer knowledge
Basic Statistical knowledge
Advanced Statistical Knowledge
Advanced IT Knowledge

Figure 18, Authors grading scale of digital tools

	Six Sigma Tools|Digital tools	Drawing Tools		Electronic Brainstorming		Internet		Spreadsheets		Statistical Software		Simulation		BI-Platform		Process Mining		Text Mining		Digital Twin		Machine Learning		Data Mining	
	Project problem, Scope and Goal			1	A									2	B										
	Process Map/SIPOC	1	A					1	A							3	C								
	VOC/Stakeholder																	3	C						
DEFINE	CTQ	1	A	1	A			1	A																
	Data Gathering							2	B	2	B													4	D
	Measurement System Analysis																			3	D				
	Descriptive Statistics							2	B	3	B			3	B	3	C							3	D
MEASURE	Capability Analysis									3	B													3	D
	Formulate Hypothesis																							2	D
	Statistical analysis							2	C	3	C													4	D
	Ishikawa diagram/RCA	1	A	1	A																				
ANALYZE	DoE									2	B	3	C							3	D	4	D	3	D
	Brainstorming			1	A																			2	D
	FMEA							1	A																
	Benchamrking					2	A																		
IMPROVE	Solution Selection Pugh Matrix							1	A																
	SPC							2	B	3	B			3	C										
CONTROL	Control Plan							1	A													4	D		

Figure 19. Authors ranking of the digital tools

The rankings are partly based on experience from working with some of the software and partly an interpretation based on issues raised in literature. For example, Schäfer et, al (2018) states "*In contrast to six sigma, the modeling phase in the data mining project is far more complex because of the huge amount of modeling techniques and algorithms*". Aggarwal (2015) also argues regarding data mining that it is difficult to create reusable and general techniques as each data mining application is unique.

As can be seen in Figure 19 most tools with high ranking in benefits also have a high threshold and often times require prior knowledge before implementing. At the same time, digital tools that does not necessitate as much prior knowledge does not add as much value compared to the more advanced tools. Due to this there is a need for a model which take both process and digital maturity into account for the framework to be applicable for different businesses. A company with low process and digital maturity is not improved by machine learning as it most likely would not be able to use the techniques efficiently. Therefore, the process maturity model in Figure 8 was reconstructed to include digital maturity and the new digital process maturity can be seen in Figure 20.

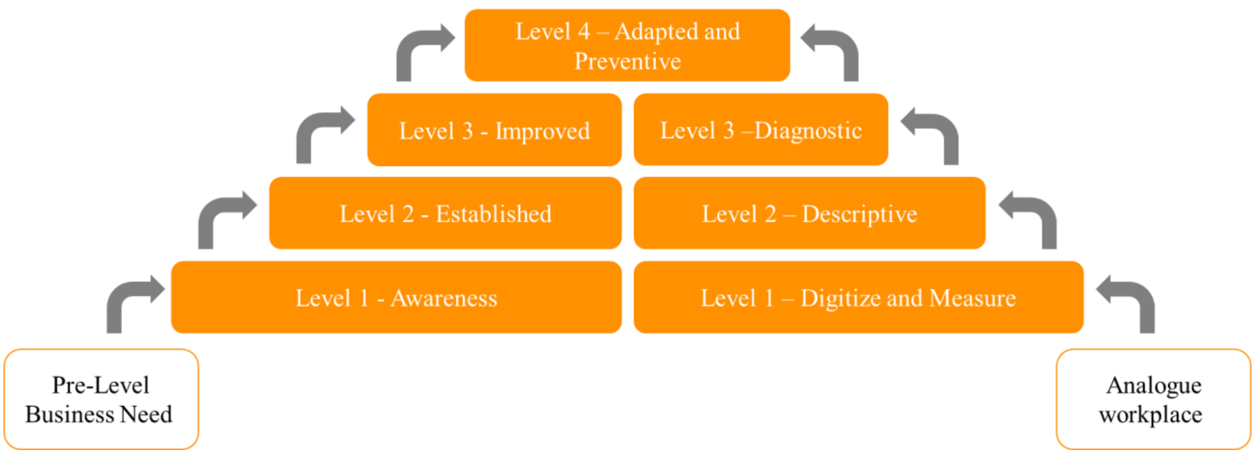

Figure 20. Digital process maturity model

The digital process maturity model aims to further establish the connection between process maturity and digital maturity. The idea is to showcase that an integration of data understanding and digital maturity can help to achieve a higher process maturity and vice versa. In the digital part of the model the levels are building blocks and as the organization become more digitally mature, the previous levels are still continuously used and act a base as a new level start to get integrated into the organization. Some parts of higher levels can be used with tools from lower levels, however this type of analysis is not efficient until an organization can utilize the more complex tools.

The first level in the model is "*Measure and Digitize*" which includes digitizing, understanding the usefulness of measurements, data, and data structure needs. It is primarily based on the findings of Sehlin, et al., (2019) as well as Chanias & Hess (2016) technological definition of digital maturity in digitizing work. While establishing a process, related measurements should be set up and the digital side therefore has a second level "*Descriptive*" which is the first level of analytics, where the organization start to visualize data to get insight into what has happened. When the organization use data visualization as support, the next step is to perform "*Diagnostic*" analytics, i.e. investigate root causes to why something happened to be able to learn and improve. The fourth and final level in the digital model is to work proactively, where predictive and prescriptive modeling is the goal, to forecast and prevent problems before they arise. For this to function, a well-defined adapted process is a pre-requisite which is why the steps are combined.

It is important to divide the digital tools between the maturity levels as the tools have vastly different thresholds and competence needs. A business with a digital maturity on level 1, "Digitize & Measure" should not try to implement machine learning or digital twins as they would not be able to reap the benefits of the tools. The digital maturity model from Figure 20 was therefore combined with the rankings from Figure 19 to showcase how digital tools can be incorporated in the Six Sigma phases. The combination between the two figures is visualized as a matrix in Figure 21.

		1. Digitize & Measure		2. Descriptive		3. Diagnostic		4. Predictive & Preventive	
	Six Sigma Tools|Digital tools	**A**	**RANK**	**B**	**RANK**	**C**	**RANK**	**D**	**RANK**
DEFINE	**Project problem, Scope and Goal**	Electronic Brainstorming	1	BI-Platform	2				
	Process Map/SIPOC	Drawing Tools	1						
		Spreadsheets	1			Process Mining	3		
	VOC/Stakeholder					Text Mining	3		
	CTQ	Drawing Tools	1						
		Spreadsheets	1						
		Electronic Brainstorming	1						
MEASURE	**Data Gathering**			Spreadsheets	2				
				Statistical Software	2				
								Data Mining	4
	Measurement System Analysis							Digital Twin	3
	Descriptive Statistics			Spreadsheets	2			Data Mining	3
				Statistical Software	3				
				BI-Platform	3				
						Process Mining	3		
	Capability Analysis			Statistical Software	3				
								Data Mining	3
ANALYZE	**Formulate Hypothesis**							Data Mining	2
	Statistical analysis					Spreadsheets	2		
						Statistical Software	3		
								Data Mining	4
	Ishikawa diagram/RCA	Drawing Tools	1						
		Electronic Brainstorming	1						
	DoE			Statistical Software	2				
						Simulation	3		
								Machine Learning	4
								Digital Twin	3
								Data Mining	3
IMPROVE	**Brainstorming**	Electronic Brainstorming	1						
								Data Mining	2
	FMEA	Spreadsheets	1						
	Benchamrking	Internet	2						
	Solution Selection Pugh Matrix	Spreadsheets	1						
CONTROL	**SPC**			Spreadsheets	2				
				Statistical Software	2				
						BI-Platform	3		
	Control Plan	Spreadsheets	1						
								Machine Learning	4

Figure 21. Merged matrix between the digital maturity model and authors ranking

The idea behind the matrix is to showcase how the digital tools can be incorporated into the Six Sigma methodology steps and how effective it would be, as well as what digital maturity is required to achieve it. To visualize that, the different digital tools were matched to all possible steps they could be incorporated in and then fitted towards a digital maturity level. The digital maturity level match was based on the threshold rankings from Figure 19 where rank A corresponds to level 1 "*Digitize & Measure*" and rank D corresponds to level 4 "*Predictive & Preventive*". The linear conversion between threshold in Figure 19 and digital maturity in Figure 20 deemed reasonable as more advanced tools requires more time, are more powerful, and build on each other. Furthermore, each tool was given a numerical rank from one to four which resembles the *added value* rank given in Figure 19 where a higher value represents a greater added value. It is also important to note that a digital tool can be applicable for different Six Sigma steps and therefore provide different added value and occur on several maturity levels.

In Figure 21, the matrix presents tools for each maturity level and the *added value* of the digital tools for each maturity level. It displays that lower digital maturity levels also have a lower rank. This indicates that simpler digital tools add less value and more advanced digital tools add more value and in order to apply the digital tools to individual business needs, a new matrix was constructed. The purpose of the new matrix is to provide a guide for selecting and structuring improvement projects based on digital maturity, data volume, complexity of the problem and whether the solution is known or not. The framework is based on Zwetsloot et.al (2018) with the additions of the digital maturity model and "Big Data" alternative in "data volume". The framework can be seen in Figure 22.

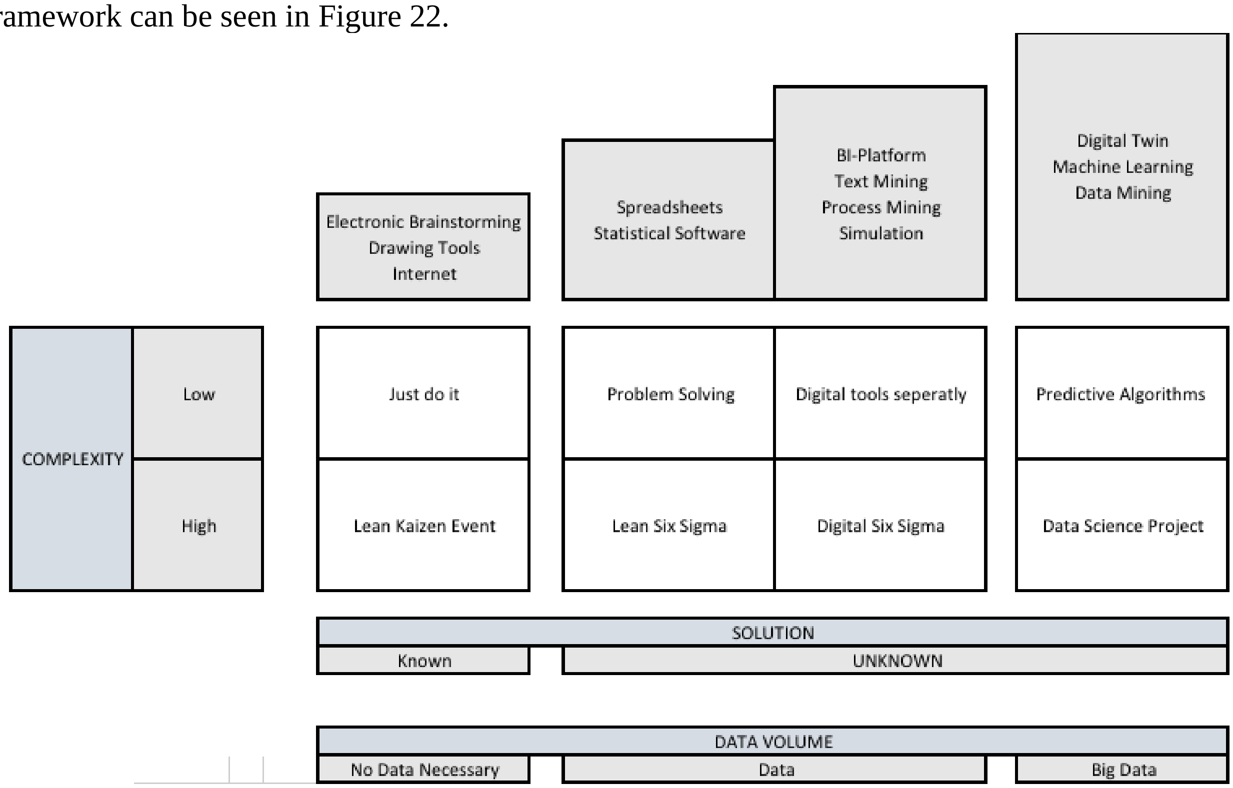

Figure 22. Matrix for improvement projects based on digital maturity and available data

The framework aims to help organizations select the correct improvement process for a specific problem and providing a structure for solving it. By evaluating their digital maturity, a business can see what different improvement process they can use. Furthermore, if a business reaches a higher digital maturity it does not imply that they should discard earlier levels improvement processes. Instead they now have two more to choose from. The different processes are meant to solve different kinds of problems.

The framework is simple to use and can be executed in four steps.

1. Look at the digital maturity of your business and assess what options are available
2. Evaluate the problem and understand whether the solution is known or unknown.
3. Evaluate the complexity of the problem, is it a low or a high complexity problem?
4. Evaluate the data volume available.

There may be cases where a business' digital maturity does not match the recommended improvement process. For example, if a problem is highly complex with an unknown solution and large amount of data the matrix will recommend "Digital Six Sigma" which requires level 3 on the digital maturity model. However, the business facing the problem has only reached level 2 and do not possess the ability to use recommended tools. In cases like this it is best to look at the axis "Data Volume" and reduce it one step. In this particular case, it would mean that the solution instead would be "Lean Six Sigma" which would be a perfectly good substitute. The best option is to reduce the "Data volume" axis one step so it matches the business's digital maturity. It will assist to solve the same type of problems, but it might require more time as the tools on lower levels have lesser capabilities.

The framework presented in Figure 11 and matrix presented in Figure 22 will be used as a basis for the empirical study to either validate or revise aspects from. Answers from the respondents will be analyzed against the theoretical foundation on which the two are based on.

5 Empirical Findings

The empirical findings are based on semi-structured interviews, held via digital platforms. This chapter contributes to the development of the final framework and supplement the theory to answer the second and third research question.

5.1 Empirical Composition

The empirical findings were gathered by conducting semi-structured interviews with four different companies and one discussion seminar with a fifth company. The interview questions were the same for all companies and connected to the three fields, Improvement Processes, (Digital) Tools, and Digital/Processes-Maturity. The areas were chosen based on the theoretical analysis to create an understanding of the interactions between a business' digital/process maturity and its improvement work as well as how digital tools can be incorporated based on the maturity levels. This was performed to confirm or revise the initial frameworks from the theoretical findings with new empirical findings. The three areas meant to complement each other as seen in Figure 23. The areas can be hard to define, a short summary will therefore follow.

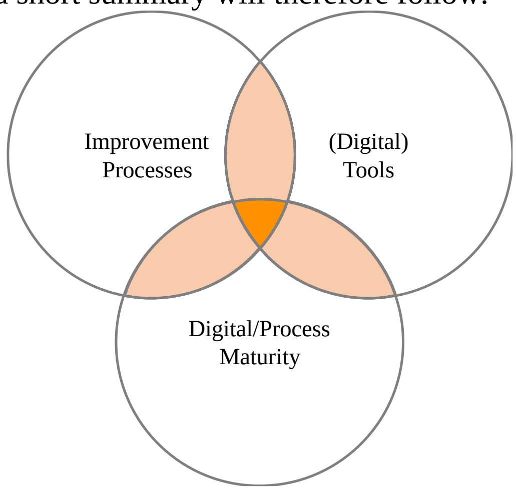

Figure 23,The three areas of empirical analysis

Improvement Processes

Improvement processes is a broad term with a vague definition and includes everything from simple "just-do-it"-activities with a known solution to complex big data projects. These problems can range from simple to extremely complex. Questions were directed to understand if business differentiate between improvement processes and if so, how they do it. Furthermore, how each business works with improvement processes and if any particular methodologies are used.

(Digital) Tools

Digital Tools as defined in this thesis is as term for describing tools and technologies that can help in improvement work. The questions asked were focused to understand which tools and techniques the companies used in their improvement work. Both analogue and digital tools were of interest to understand possible correlations to the companies digital and process maturities.

Digital/Process-Maturity

Process maturity in this thesis is defined as described by Cronemyr & Danielsson (2013). This is used as one of the case companies has used the exact model to develop their business, the model is also used by the thesis constituents Propia. Digital maturity is the authors attempt to create an extension to the existing model but base it on how businesses work with data. The questions asked were directed to understand how different maturity levels relates to digital tools and techniques used in improvement work.

5.2 Case Companies

The following section will describe the case companies and a summary of the interviews performed divided by the three areas.

5.2.1 Tekniska Verken

People in different roles from two separate business areas were interviewed. The two areas were biofuel-based energy department and power grid department. Throughout the company it is now a comprehensive focus on implementing a BI software package. The business areas have taken somewhat different approaches in their respective implementation of BI software where one function uses a small-scale implementation while the other wait for the full-scale implementation.

5.2.1.1 Tekniska Verken Biofuel-Based Energy (BBE)

From the business area of biofuel-based energy people in roles connected to business development and process management were interviewed. The summary of each subject is a compilation of issues raised by all respondents which were similar as they work in the same business area.

Improvement Processes

The foundation of the improvement work is built upon the work with processes. All employees are encouraged to submit improvement suggestions, there is a report system to submit the suggestions but there is no limitation in what way concerns are raised, as long as they are. The suggestions are then either immediately implemented if there is a possibility for a quick alteration in the process, otherwise it is brought up at a recurrent process meeting. There, decisions on which improvements to move forward with are made based on prioritization in the meetings, sometimes based on gut feeling and to some extent weighing time and cost against benefits. Larger improvements are assigned to a project group to take responsibility of and move forward with. The work with smaller continuous improvements is considered part of the process whereas larger changes need to be lifted out of the process work. The idea of getting the small fixes and improvements continuously is to enable more free time to work with larger improvement projects. However, they have noticed that some problems reoccur after implementing changes.

Often there is a lack of follow-up on improvements and changes which results in that the achieved effects of the improvement work are not visible. This is a concern as it leads to lower motivation to work with improvements and problems reoccurring. A current focus is therefore to start using a standardized template for improvement work throughout the business area. A problem raised regarding follow up is how it should be measured effectively to be able to follow up.

(Digital) Tools

Different support systems are used for carrying out different processes and for the administration of operations. Common tools throughout the organization is a modelling tool to visualize and establish processes with their respective activities and documentation. This process map serves as a guide for how the process should be executed and what activities are linked. Data is generally analyzed through different Excel files and used for following up and planning. Excel is however not the tool wanted to be used forward but rather support systems and a BI platform as it should provide a more comprehensive picture and reach more people.

The main advantage of implementing a BI platform is to get the collected data gathered in one place to enable and simplify follow up on a higher level and to see the effects of the improvement work. This should also automate a lot of the work currently performed in Excel files. The unification of data is also intended to be used for more specific analyses on the process level by drilling down in the data. It is challenging that the process teams are not familiar with the tool yet and the data team does not understand the processes where the data is collected.

Digital/Process-Maturity

Processes are in different stages of maturity as it is time consuming to establish and implement the processes. To start with improvement work, a process needs to be mapped and a process owner and group assigned. Even though the roadmap for reaching higher levels is known, it requires constant work and is time consuming. An understanding of what benefits the way of working yields is sometimes lacking and some resistance exist. Almost all work is digitized and carried out through support systems but supplemented with various different Excel files. With the large amounts of existing systems, it can emerge resistance to new systems being introduced and therefore new systems should only be introduced to those who need it. The scattered systems and files are not always unified as they can utilize different units for measurements or different ways to identify a customer.

5.2.1.2 Tekniska Verken Power Grids (NÄT)

One interview was conducted at the area of power grids with two respondents simultaneously, the interviewees were a business developer and a manager for business development.

Improvement Processes

All improvement work is based around a process methodology which aims to describe processes from the employee's point of view in order to engage the whole organization. The improvement work proceeds from the concept that employees are encouraged to raise improvement proposals regarding processes. The proposals are then discussed and prioritized at a monthly meeting together with process owners, process leaders, and the employees who raised the suggestions. The improvements are ranked according to criticality, simplicity in implementation, gut-feeling, and occasionally economic feasibility. Smaller changes are managed by process leaders while bigger changes are realized in project form with an assigned process group. The perception is that a process group is well suited to drive continuous improvements within their process whereas the temporary nature of a project fits well with larger alterations.

The vast amount of improvements are smaller changes which can be solved by quick fixes for which there is no specified methodology. There has been Kaizen initiatives in the past, but the improvement work has shifted into utilizing the process methodology instead. There is drawback with this as these improvements heavily rely on the process owners to realize them. There is an apprehension that lack of commitment will result in fewer implemented solutions which might impact the submission rate of future improvements. The larger problems are usually solved in projects according to an old project model. The model helps clarify deliverables, scope, and goals. The problem solving is performed by deconstructing and discussing possible improvements within the project group. The improvement work is rarely data driven.

(Digital) Tools

The most used digital tool is Microsoft Excel, there are however ongoing attempts to start implementing BI. The ambition is to transfer the Excel files into a structured data lake which can provide Power BI with data. They have begun initial small-scale trials to incorporate a few Excel sheets into Power BI to get an understanding of the software. The early attempts have already paid dividends as they are able to produce detailed cross-function reports. They have for example produced a BI-report which exposed identical oil leakages at ten similar machines. This discovery prompted them to change the remaining ten before any leakage occurred saving them time and money.

The analysis is still performed by a human looking at the same data they had in Excel but now it is displayed in colorful graphs with structure. They raised concern that bad data quality might lead to wrongful decisions as the data is the same as before but presented in professional looking graphs. People may therefore assume it is correct simply because it looks nice and do not critically approach it. Furthermore, it is still decisions based on visualization and gut feeling, there is no statistical significance to reinforce them.

It has been hard to find business cases that motivates large investments in new digital tools. A comparing example of power theft in Brazil by bypassing electricity meters where AI-investment had been justified was that theft accounted for approximately 35% of all power consumption, whereas theft accounts for minimal part in Sweden. However, with being able to generate detailed reports between functions they have seen that they can use data in ways now that they previously needed to buy separate packages from a third party to use. The reports created in house can also be tailored more to the specific needs and functions actually wanted.

Digital/Process-Maturity

Process maturity is varying a lot through the organization, with some processes in the improvement stage and some yet to be established. There are limited available time and resources to work with more than a couple of processes at once, and therefore it takes time to implement new processes. They often know what is missing in order to gain a higher process maturity, e.g. lack of process measurement. The work is generally digitized and process improvements are closely connected to improvement in support systems. Data has started to be shared across departments and thereby used for new purposes with the implementation of BI and the hope is to further develop in this area to solve more problems with data analysis. However, improvement work is rarely data driven, certain statistics are sometimes looked at when possible.

5.2.2 TitanX

TitanX is a global supplier of powertrain cooling solutions with plants around the globe. The company has a production plant in Linköping which was object for the interview. TitanX has knowledge of structured improvement work in the form of LEAN. They have also basic understanding of Six Sigma as they have acted as a case company for student lead Six Sigma projects. The interview was held virtually with TitanX quality manager. TitanX are not working with Propia's process maturity model. Therefore, the questions regarding process maturity were remodeled towards digital maturity as described in this report.

Improvement Processes

The focus of improvement work lies in waste reduction within production and due to capacity limits, it is important that the equipment is constantly producing. This often mean that quick fixes are necessary, rather than executing a thorough investigation. Efforts are being put in to work more proactively with finding underlying problem causes with improvement structure such as DMAIC and PDCA which are being used to some extent. The methods are primarily used for larger complex problems with unknown root causes. They have one Six Sigma master black belt and several green belts that possess the required statistical knowledge to conduct Six Sigma projects, however the improvement work is seldom structured as quick fixes need to be prioritized.

There are three big obstacles when working with improvements, the first being time limitations. It is a result from the production need and capacity limits which leaves no time for systematic improvement work but instead concentrating on short-term quick fixes to maximize production uptime. The second obstacle is endurance in the follow up on improvement projects. A project is deemed finished when an improvement is implemented. They have a difficult time measuring and following through to see if the implementation was a success or failure and thereby have difficulties sustaining improvements. The third obstacle and what they believe is the biggest challenge is that there is no standardized way of working. As of now a lot of projects are initiated but there is no clear definition of the problems, various approaches being used and no follow up. They want a methodology that could help structure improvement work for diverse type of problems. They would greatly benefit from a guide that could recommend a structured approach depending on the problem which then suggests possible tools and techniques to solve the problem.

(Digital) Tools

TitanX is working with several different digital tools. The first digital tools are when gathering data. Data collection is performed from two different systems, the first being the ERP, Enterprise Resource Planning System, which provides a holistic view of the production flow. However, specific data points to troubleshoot problems and investigate improvements, cannot be found through the ERP as it only measures logistics measurements such as throughput time, inventory levels and lead times. These measurements can be used to find anomalies or possible improvement opportunities which then can be investigated further. The second data collection is automatic and continuously gather detailed data from processes into an Excel-sheet. This data is used to understand the problem and is used for statistical calculations to find root causes. Sometimes the data collected is not sufficient to solve the problem in which case a manual data gathering needs to be conducted. The measurements are performed by operators which is then typed into an Excel-sheet.

Microsoft Excel is the primary tool used for visualizing and analyzing data. Minitab is sometimes used for more complex problems which requires statistical analysis. They use Pareto charts in Excel to visualize weekly scrap quantity and from that derive where to focus the improvement work and quick fixes. For more complex analysis fishbone diagrams, brainstorming and statistical process control has been used. These tools are sporadically used when they have time.

They see the potential with integration of digital tools into structured improvement work. They believe a BI-platform could yield benefits by providing traceability and a holistic view of data which is now scattered throughout several Excel-sheets. They have however just integrated a new ERP system to replace the old one which leaves no economical leeway for more investments.

Digital/Process-Maturity

TitanX has a lot to develop regarding digital maturity. They are using numerous disorderly Excel files constructed for specific needs. They see potential for development if these can be combined into one spreadsheet or integrated into a new software. Furthermore, they still use a lot of paper as it is easier to bring a sheet of paper than a computer into the production unit when troubleshooting. They do however acknowledge the drawbacks that the information might not be digitized and if it is the lack of standardization might entail faulty translations which renders it useless.

5.2.3 Sandvik

The interviewee is working with continuous improvements in one of the production plants at a subsidiary to the global engineering group Sandvik Group and is also Six Sigma Black Belt. Sandvik invest a lot in new technology and digitalization throughout their companies to improve customer offer as well as internal processes.

Improvement Processes

They are working with improvement process in many different ways, primarily utilizing a bottom up approach as the operators are encouraged to raise improvements and problems in the production which are then discussed in improvement forums. Furthermore, they are also using a top-down approach for larger strategic improvements. These are more often based on data in the form of key performance indicators as opposed to intuition from operators. It is a combination of short-term and long-term improvements. The short-term work is primarily performed through continuous improvements, while the long-term improvements include strategic decisions based on trends and aggregated data.

The improvements methods used are mainly A3 and PDCA which are both structured data driven methodologies. The DMAIC structure is not commonly used even though they can see the potential as it is more detailed than the other two. However, the method requires more training in addition to already having two functioning methods.

(Digital) Tools

They have come a long way working with digital tools with the core in their production being an advanced OEE (Overall equipment effectiveness) system with embedded sensors which enables data driven production. It feeds vast amounts of data into databases which then can be analyzed in a BI-platform. The enormous amounts of data support data-driven decision making and improvement work as the data can be accessed by many throughout the plant. The BI-platform is primarily used to visualize the vast amounts of data collected to monitor trends and patterns closely. The perception was however that the software might not be fitted to enable a thorough root cause analysis as Six Sigma needs customized and very specific data which the standardized reports in Business Intelligence software cannot accomplish. He could however see the benefits of such systems when trying to define problems, initially measure them and particularly in the follow up of effects of projects.

There are also projects trying to utilize techniques to enable machine to machine communication and automatic feedback based on data. These are however still in the development stage and therefore confidential. There have also been development projects of mobile applications in order to simplify and increase the submission ratio of improvement proposals. They also have an ongoing project with incorporating multivariate analysis in order to predict and steer production processes in real time.

Digital/Process-Maturity

The plant is very digitally mature, mostly because of the sophisticated OEE system equipped with IoT sensors collecting very detailed data. As they are utilizing data in their everyday work, both for short-term improvements as well as long-term strategic decisions, they are very cautious and work a lot with securing data quality. Data is being used in a descriptive, diagnostic, and somewhat predictive manner in order to base their decision on facts. A lot of digitalization efforts are also put in towards enhancing the customer offer.

5.2.4 Ericsson

Ericsson is a Swedish telecommunication company and one of the leading producers and provider of mobile communication networks. Their main focus as of today is the development of the 5G networks, IoT connectivity, and automation technology. The interview was conducted with two employees at Ericsson, one senior process manager and one business performance & improvement manager who is also a Six Sigma master black belt.

Improvement Processes

They are working rigorously with improvements at Ericsson. It is hard to generalize the improvement work due to its size and global presence, but the common ground is to work data driven. They use the DMAIC methodology, also known as Six Sigma approach, for all kinds of problems. It is used due to its structured approach in defining the problem, finding the root causes, and following up on the implementations. They have an internal Six Sigma training program where the methodology and different tools are taught in gradually increasing difficulties. The training is conducted to provide knowledge regarding continuous improvements and problem-solving methodology as it is each individual’s responsibility to raise proposals to improve the business.

Even though the DMAIC structure is used for all problems the specific procedure may differ between different problems as it is not always profitable to conduct deep analysis. Smaller problems do not need equally extensive phases or the same tools that larger problems require. For example, a DMAIC-light is used for continuous improvements which utilizes the same structure but on a basic level with some simple tools. The structure helps enables everyone in the company to improve continuously as they all know the procedure of the improvement methodology.

(Digital) Tools

They utilize several digital tools to analyze their business. They have a designated group for data and business analytics which purpose is to facilitate data platforms to incorporate scattered data into data warehouses and provide analyses of this data. Ericsson has a long history with digital tools and implementation of BI, which is widely used within the business today, approximately 10 years ago. It is used to measure key performance indicators, check for deviations, and obtaining structured reports. Their latest investment is the incorporation of process mining into the business. It is used to analyze internal processes by process adherence and process compliance which measures if the workflow follows the supposed one or if it deviates. It can also be used to detect bottle necks, measuring lead-times, finding and optimizing processes and the ambition is that it should have a place in all of the phases of DMAIC. Process mining is also used to find automation opportunities. In addition to BI and process mining, statistical software is also used as it is complementary rather than supplementary to process mining. They are also currently working to integrate advanced machine learning, data mining, and artificial intelligence in their solutions. Customers are requesting it as the automation market is increasing.

Digital/Process-Maturity

Ericsson are very digitally mature with a high process maturity and experience from process-oriented work, with their old process maps now beginning to be analyzed against computer systems with process mining. They have a substantial analytics teams for building data capability and support in analyzing the business. They conduct improvement work training courses to teach employees the same structured methodology. Furthermore, they use a data driven approach for both advanced problem solving as well as for continuous improvements.

5.2.5 Propia

Propia is a consultancy firm specializing in process management, change management and business development and are also the constituents for the thesis. A presentation of the progression of the thesis was held followed by an open discussion to receive input and feedback regarding the theoretical analysis, empirical findings, and what should be included in a final result. It emerged that the final result should include distinct activities connected to follow-up and implementation of solutions and a flow chart structure were favored because of its pedagogic nature.

6 Analysis

In this chapter the frame of reference and results of the theoretical analysis are combined with the empirical findings for further analysis. Insights from the analysis are used to further develop and revise the final framework.

6.1 Improvement processes

Consistent through all case companies is that they work differently with different types of problems and improvements. All prefer to have some kind of bottom up continuous improvements structure in place to catch smaller problems and improvement opportunities. There are some differences in the approaches taken and to what extent it is data driven. How much digital tools are included are dependent on what scale data are collected and how mature the companies are.

The empirical findings in both business areas of Tekniska Verken show strong focus in current work with establishing the TQM cornerstone *focus on processes*, to lay the foundation of structured improvement work. Therefore, there is a lot of small improvement projects, kaizen like activities, that does not incorporate data to solve problems. These would be categorized at level 1 in Figure 22, as they are either implemented immediately, *Just do it*, or solved through workshop events, *Lean Kaizen event*. Troubles in tracking the effects of the implemented changes and in using the correct measurements for the processes are an obstacle in moving further on the maturity scale. Even though there is “no data necessary” to solve the problems, data is still needed to follow up that it was an actual improvement. Data analysis cannot be incorporated in the improvement work if there are no relevant measures to analyze. Improvements projects are mostly initiated based of measurements and key performance indicators. Isolated measurements are studied rather than a thorough data analysis and then a current and future state is sketched before using problem solving and workshops. Therefore Level 2 approaches is somewhat used as well. Even though Tekniska Verken are working with improvements there is a consensus that the follow up is insufficient and they are not sure if they are able to sustain the improvements. This occurrence is prevalent and a large obstacle according to all respondents and is in line with common limitations of change management. They fear that a lack of follow up might impair the quality culture, the employees will stop raising improvements as they see no change from previous proposals. Bergman & Klefsjö (2010) argues that corporate culture is vital to a business’s well-being and a good culture will lead to increased profits and improved results.

At TitanX almost all problem solving is data driven to some degree, starting from pareto charts. However, they do not generate a lot of data as they have mostly manual data collection and the problems cannot be considered data rich. Furthermore, they do not utilize any specific methodology but rather use tools and techniques which they find suitable for the specific problem. Approaches utilized are thereby on the lower end on the *data volume* axis in Figure 22, with most of the work positioned on level 2.

Both Ericsson and Sandvik utilizes structured data-driven methodologies for improvement work. Ericsson use DMAIC while Sandvik uses A3 and PDCA. The structure of the methodologies is used for all types of improvement and problems, the difference being in how deep they go. All problems do not need the same analysis, but the same structure can be used regardless of problem.

Both companies heavily rely on their employees to raise improvement proposals and emphasize on the need to continuously improve which is one of the TQM cornerstones, *let everyone be committed*. Furthermore, they complement their bottom-up approach with a data-driven top-down approach as they base their strategic decisions on facts which is another of the TQM cornerstones. To be able to base decision on facts they collect and analyze vast amounts of data with complex tools in structured ways which would put them on the higher end of the *data volume* axis in Figure 22 and level 3 to 4 in the maturity model.

Even though the specifics in the methodologies differ, the common denominator is that they are theoretically based improvement structures for finding root causes. The theories behind the methods are quite simple and easy to learn yet powerful when searching for root causes and continuous improvements. Despite the simplicity and potency behind the methods it is only used by the two global corporations Ericsson and Sandvik. TitanX has tried to implement PDCA and DMAIC but has not fully succeeded as they still mostly use individual tools to solve problems without an underlying structure. Tekniska Verken has instead opted to use a process management thinking to perfect the processes. They tackle smaller problems in workshops and larger in project form with free reins and little to no underlying structure. Through their process orientation they are currently trying to establish a standardized improvement process that should provide structure.

The difference between the four respondents is that Ericsson and Sandvik, two global world leading actors, are continuously working with the same structured data driven approaches to solve problems compared to the other two which are utilizing different methods, tools, and techniques for different problems. The two big corporations are also dealing with exceptionally greater quantities of data and have far more people involved. While it is true that the capital difference is massive it should not impact the structure used for improvement processes. The reason behind this distinction should not only be derived to capital means but also include the parameter how long the companies have worked with the methodologies. Both Ericsson and Sandvik have worked with the same methods for a long time and the methods are today an integrated part of their respective businesses. Employees are schooled in the structures and are expected to understand and use them. It may sound simple, but time is a scarce resource and it takes time to build well-functioning improvement processes. Furthermore, it seems like the choosing of method is less important than time spent using it as long as it is a widely recognized data driven methodology. A successful way for improvement work is to settle on a method and continue using that and it will pay dividends over time.

6.2 Digital Tools

All interviews when BI was discussed the need to take backend structuring of data into account was raised. It is an important foundation to further incorporate digital tools to improve and use more complex methods to extract value from the data. They also further emphasized on the importance of good data quality. A few of the interviewees talked about the dangers of making decisions on faulty data which might be more malicious than basing decisions on gut-feeling as they then are in belief that the decisions must be correct.

When collecting vast amounts of data, it can however lead much closer to the root cause and redefine larger problems to smaller as a result, e.g. the OEE system at Sandvik that can pinpoint very specific details in circumstances surrounding a defect. The widespread BI reports used around the plant is however centered on simple visualizations such as pareto charts but with more details to sort by than TitanX's data driven improvement and problem solving primarily centered on pareto charts in excel files with limited data. TitanX raised the issue that the new tools probably would not be as efficient for them as they do not deal with big data in this context. The difference between the two approaches and tools are therefore the amount of data easily handled, real time or not, and proliferation of the knowledge through a standardized report rather than single spreadsheet files.

In single cases at Tekniska Verken the implementation of a BI-platform has enabled a quicker, more in-depth data analysis by consolidating previously isolated data sources into one shared report. The example of a quickly discovered oil leakage by simply having access to the data via the use of a BI platform correspond to the category *digital tools separately* on level 3 with low complexity. If smaller problems go undetected it might accumulate into greater problems, this can be averted by continuously analyzing large datasets. The interviewee from Sandvik shares the view of Duarte (2017) that BI software packages provide a localized view of many variables but is lacking when it comes to providing an end-to-end view of a process. BI platforms are however very effective to locate problem areas to initiate a project and when standardizing solutions into the process and follow up on them as part of the Control phase of DMAIC. Furthermore, there are differences in ways of using BI. The first being as an advanced Excel-supplement which can gather scattered data files into one place in order to provide a common platform and a holistic view. The drawback is that there will be a delay as the Excel-files needs to be updated manually which corresponds to level 2 in the maturity model. The second way of using BI is as a real-time detector with live data to monitor deviations, initial diagnostic analysis, and used for follow-up which corresponds to level 3 in the maturity model. In order for this to work there needs to be a data warehouse in place which gathers and cleans data which is then fed into the BI-platform.

Ericsson's ambition is to integrate process mining into the entire DMAIC structure with the most usefulness believed to be in the Measure and Analyze phases. To achieve a successful integration, their process mining tool are becoming an integral part of their Six Sigma training. This coincides with a suggestion of changed Black Belt curriculum laid forward by (Zwetsloot, et al., 2018) to incorporate data science topics in Black Belt training. Statistical software and other separate analysis tools still have a part to play and process mining should be used in collaboration with the other tools rather than replace them completely, to extract even more information on the process and problems. Process mining is also used as the first step in automation discovery which Geyer-Klingeberg, et al, (2018) argues that process mining is a great tool for, to show whether a process is standardized enough for automation or otherwise to find root causes for the variations.

Sandvik's efforts to perform predictive analytics with multivariate analysis to monitor and steer processes coincides with the arguments presented by Zwetsloot, et al. (2018) and Fahey, et al (2020) that traditional methods are not as effective when dealing with large datasets characterized by the "V:s" of big data also known as volume, variety, veracity, velocity, . Development projects to enable machine to machine communication with actions recommended by the machine fall under prescriptive analytics. These types of projects are however somewhat detached from the continuous improvement work and if fitted into a DMAIC structure, would probably be solution sub-projects as a part of the Improve phase. Initiatives like this are also dependent on a large-scale real-time collection of data and therefore requires investments in installed hardware to utilize.

Case studies brought up in section 3.5 utilize many new and sophisticated technologies, yet there is still a need for qualitative tools such as Ishikawa diagrams and 5 whys when performing root cause analysis. However, the logical flow of DMAIC might be altered slightly, e.g. Ghosh & Maiti (2014) use Ishikawa diagrams in the measure-phase before applying data mining algorithms which they use to avoid physical experiments to conduct root cause analysis. Not even the most advanced tools, for example machine learning can replace all tools in an improvement project. Digital and quality tools together primarily help to narrow the scope and number of variables needed to investigate as when the number of variables grow large enough, spurious correlations could outnumber meaningful ones (Calude & Longo, 2016). Fahey et, al. (2020) bring up that root causes might not be found in the data and need a Cause and Effect session to be found. The pace of root cause identification was however raised as a significant benefit. The empirical studies confirm that qualitative tools are used to complement the newer digital tools while conducting improvement projects as the complete structure of known improvement processes such as DMAIC, A3 and PDCA are utilized.

6.3 Digital/Process-Maturity

All case companies can be seen as digitally mature in the technological definition by Chanias & Hess (2016), regarding the extent to which business tasks are handled by IT-systems. The managerial aspect of digital maturity, the status of an organization’s digital transformation differs more between the companies. According to Kane (2017), digital transformation is adopting business processes and practices that help to compete effectively in an increasingly digital world. Sandvik and Ericsson have built both process and digital maturity as well as digital capability among employees over a long period of time. They have also put in large efforts to secure data quality and outlined opportunities for using the data before implementing the new technology. They look at what they want to achieve and then gather the data to do it rather than gathering a lot of data and then trying to find opportunities to use it. The digital tools and techniques they use themselves are also something they want to incorporate in their customer offer to enhance it, thus incorporating the three principles for successful big data analysis, *data quality*, *methods of data analysis*, and *customer perspective* described by (Su, 2019). For example, Ericsson noted that a higher digital maturity of their organization significantly helped when introducing process mining compared to when they much earlier implemented BI. They thought that one is probably not easier than the other but rather process mining is implemented once you are mature in this context and therefore it was perceived easier as the logic to implement correctly was already known in the organization.

Examples from Tekniska Verken show tendencies towards maturity in the managerial aspect as the technology enable a change in operational processes such as BI removing the need to purchase add-ins from suppliers as well as enable closer collaboration between departments. By gathering scattered information into the BI-platform everyone can access the same information at any time. Furthermore, Tekniska Verken has begun to implement process thinking into their business. The process maturity levels differ between processes and departments but there is a drive to continue working in order to obtain a higher maturity. There is a belief that continuous process improvement will lead to a more effective business and the possibility to implement more digital tools to work data driven.

TitanX has well defined workflow but does not explicitly work with processes. This may be derived to operating in the manufacturing industry where workflow and products are mostly standardized with well-defined activities. Both TitanX and Sandvik are manufacturing companies and are utilizing a workflow closely related to process orientation even though they may not be aware of it. Well-defined activities and process maps are of great advantage when collecting data by showcasing which IT-system are connected to which activity and thereby help to integrate new digital tools. Distinct delimitations will help single out activities to get correct measurements and distinguish data between processes. A high process maturity can help to raise the digital maturity and a higher digital maturity with integrated tools such as BI and process mining can help to define processes and thereby raise process maturity levels. Ericsson explained that they are currently using process mining for this purpose by comparing their previous process maps to the digital footprint activities have. The two maturity levels, process and digital, impact each other and can help to raise each other as long as the business utilizes its strengths.

6.4 Key takeaways for revision of framework

In a structural context, the findings point to a need to separate the final framework based primarily on what is included in a continuous improvement structure and what is dealt with as projects and then gradually increasing data volume and tools included in the structures. Using a structure builds a quality culture and by extension a maturity within the organization, exactly which structure or improvement process being used is not the crucial part. As presented in section 3.1, many of the improvement structures are similar and consist of almost the same building blocks. An interesting finding was that companies use the same improvement structure for all problems but with different depth analysis. A vital distinction within the expression improvement processes is "continuous improvements" versus "problem solving". The first describes a constant effort to improve a process or business by introducing small changes. These changes often come from employees close to the given process who can discover the small possible improvements. The second expression, problem solving, is an umbrella term for different methodologies used for finding root causes to unknown problems within the business. Therefore, a revision of Figure 22 was made in order to match the findings. The revised figure can be seen in Figure 24.

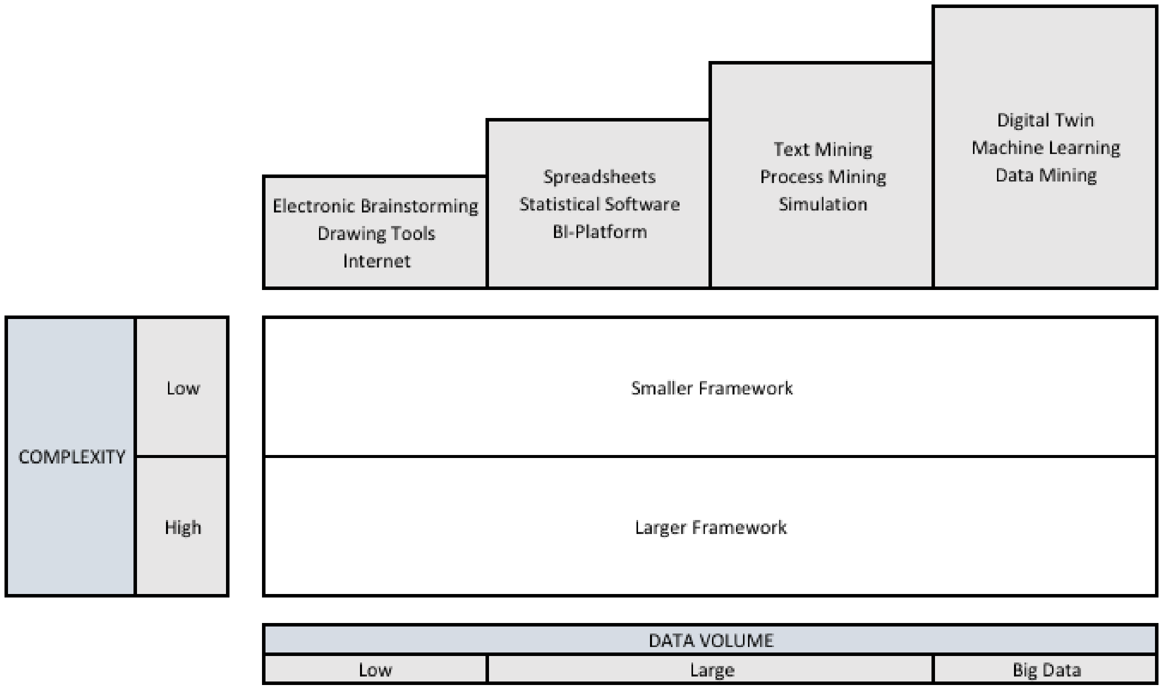

Figure 24, Revised Matrix for improvement projects based on digital maturity and available data

It seems like the foundation for building a high maturity is a familiar improvement structure which is used for all kind of problem solving. As mentioned before the specific methodology is not as important as the structure itself. This structure should then be iterated within the business to solve problems and improve processes. This will in turn raise the process maturity which can help the digital maturity to increase which in turn enables new digital tools to be incorporated into the existing improvement structure. It is an iterative process which help businesses to solve problems, improve continuously and raising their digital and process maturity. The process is illustrated in Figure 25.

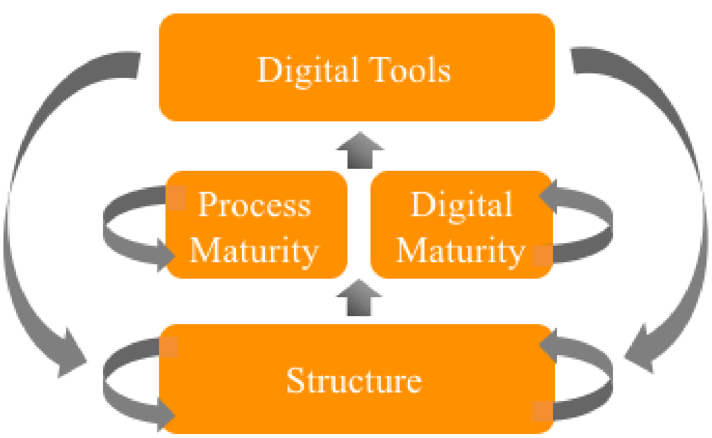

Figure 25. Visualization of how to build maturity

Several respondents faced a challenge defining the correct measurements and following up on them. This is one of the most essential parts when working with improvements and problem solving. New output values need to be compared to old output values in order to ascertain that a change was an improvement and not a deterioration.

Another key take-away connected to the maturity model and incorporation of tools is that newer digital tools should not replace older tools but be used together to speed up projects while yielding more significant insights. BI provide easy access to data and when a maturity and capability has been built around the tool it can be used to monitor trends and patterns closely in real time. It plays its biggest role in the follow up work and when detecting issues. Process mining is instead a larger part in performing more in-depth analysis centered on the process and completely integrated in more comprehensive improvement projects. Respondents said BI is usually implemented before process mining which prompted a degradation of BI from level 3 to level 2 in the digital maturity model. Furthermore, tools like data mining and machine learning can be incredibly helpful in larger projects but to find root causes a simple tool as Ishikawa diagram paired with 5whys can be desirable.

7 Result - Final Framework

The results are presented in the form of two improvement structures, one for larger problem solving and improvement opportunities and a smaller continuous improvements structure. The frameworks are presented as flow charts with yes and no questions to guide the user towards possible tools.

The final framework consists of two different models for integrating digital tools in improvement work. The first model describes an approach for larger problems with a more in-depth analysis while the second one describes a fast-paced improvement work structure. The improvement structures in the two frameworks utilize the same logical structure while the comprehensiveness of included activities and examples of tools differ.

The frameworks incorporate the levels presented in the maturity model extended to include digital maturity by matching the tools and techniques according to the digital maturity levels seen in Figure 26. This was done to present that the use of digital tools in improvement processes are ongoing and become more efficient as the organization matures. The digital tools and techniques positioning on the various levels are ranked according to difficulty, generalizability in use cases, depth of analysis and connection to process maturity. It is important to understand that the frameworks are used as a mean to display possible digital tools and techniques that can be used to facilitate improvements based on their current process and digital maturity. The thought behind it is to reinforce mature companies with appropriate digital tools not as a mean to show tools that can be used in order to reach a higher maturity. This will instead be counterproductive as the higher-level tools will be too advanced which will result in the business not being able to handle them and they will not add any value. The authors have delimitated from an exact evaluation of an organization's digital maturity level, but indications can be drawn from the process maturity to match its digital maturity. Other parameters that could be included in the evaluation is to what extent improvement work is data driven and how extensive and automated data collection procedures are.

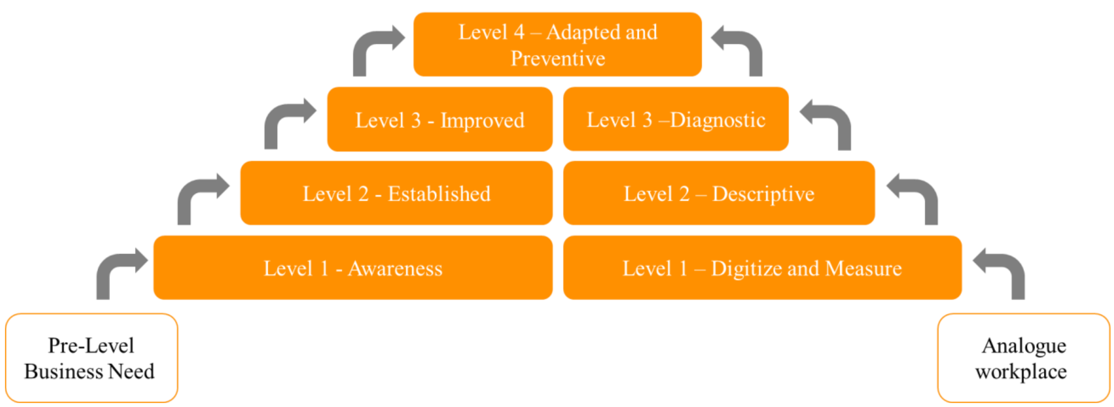

Figure 26. Combined process and digital maturity model

The larger framework is loosely based on the DMAIC structure with defining, measuring, analyzing, improving, and following up the problem. It is an attempt to simplify a problem-solving process by providing a structured methodology while incorporating digital tools based on a business maturity. The second structure is intended for small problems or improvement opportunities that occur as suggestions raised by people in the process. As the empirical findings suggested, improvements do not need the same in-depth analysis as a complex problem needs. Therefore, a second framework is presented which is intended to be used as a fast-paced way of improving a business. The structure of the framework is the same as the bigger framework but excludes the exhaustive analyze parts and has less in-depth questions to be answered as it is not needed in smaller improvement projects.

Another field of application of the smaller framework is to act as a learning platform to create familiarity of the structure and as a basis for creating a quality culture. By using the smaller framework employees will understand the structure used for larger projects but more importantly will be able to carry out continuous improvements by themselves. A quality culture is built on the foundation of letting everyone be committed and by encouraging employees to use the smaller structure a quality culture can begin to form.

As described earlier, problems and improvement opportunities can appear in various ways with common examples being suggestions from people working in a process, deviations while monitoring the process or customer feedback. Typical projects that should utilize the larger structure is for example reoccurring problems or process redesigns. The smaller version is instead suitable for quick fixes and small process alterations where trial-and-error can be suitable.

7.1 Larger Framework for Problem Solving

In Figure 27 an overview of the larger problem-solving structure is visualized. The left part of the figure is a flow chart structure which guides the user by asking questions and depending on the answers given, different actions should be taken.

The right part of the figure consists of rows of four interconnected boxes which are linked to a specific step in the flow chart. These boxes in the rows contain digital tools which in turn are related to a certain maturity level, the levels on top of the proposed activities and tools represents the digital maturity levels. Depending on a business digital maturity different tools can be used to facilitate the particular phase in the flow chart. The boxes in the flow chart are then presented more thoroughly together with examples of the digital tools below the framework.

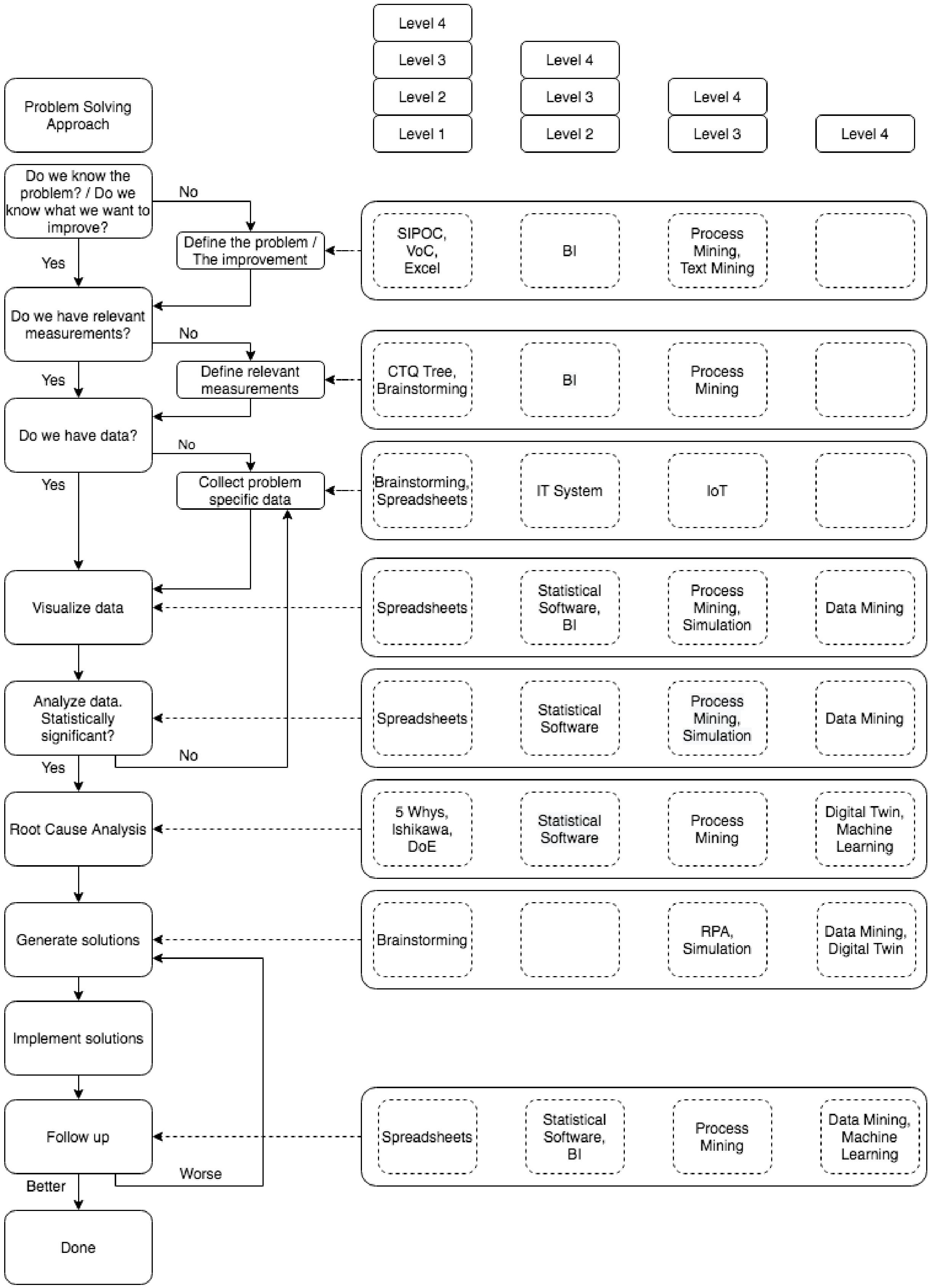

Figure 27. Problem-solving structure

Do we know the problem or what we want to improve?

The first part is to make sure the scope of the problem and improvement goal is understood and clearly defined. This includes connecting it to a process, internal or external customers and constructing a project team. This part is important to know whether goals are met, to motivate costs of conducting the project, to understand how a process works and what the customer expects.

Questions to answer

- What is the scope of the problem?
- What process is the problem connected to?
- Who is the customer, internal/external?
- Who should be part of the project team?

Define the problem / improvement

This step is performed to answer the questions from the previous step if they were not fulfilled. The tools that can be used to define the problem for each maturity level are explained below.

Level 1

Setting the scope of project involves problem and target statement and putting together a team to drive the project. If there is some data available to support the activities, spreadsheets are often used, for example Excel.

Connecting the problem to a process could be easy if the organization have established and mapped out their processes. If they do not have established processes, a process map or a SIPOC should be constructed, often done by interviews and workshops, then put into a modelling tool for digitization and simpler overview. Through the process map/SIPOC it can also be seen which departments or roles that have knowledge about the process and can contribute to the team.

To gather the Voice of the Customer several methods are possible, surveys or feedback from external customers and interviews or workshops for internal customers. It should then be compiled which the Kano model is used for.

Level 2

BI-Platforms can highlight trends or show deviations to help with the problem statement as well as narrow the scope and help quantify the problem.

Level 3 and 4

Process mining step discovery can be used to extract a process map based on an event log from internal systems, it can provide initial insight to problems if compared to existing process maps from modelling tools. It can also discover problems such as bottlenecks or non-value adding activities.

Through text mining a more extensive VOC can be collected by for example gathering customer reviews and classify them into groups.

Do we have relevant measurements?

This is one of the most important parts in the framework as the only way of knowing whether a change was an improvement is to measure the output before and after. It is therefore important to have relevant measurements for the specific project.

Questions to answer

- Are the measurements for the process relevant?
- Do they reflect what is expected of the process?
- Can they be affected by an improved process?

Define relevant measurements

It is vital to define the correct measurements for the process in order to be able to follow up the result of the improvement. Therefore, it requires personnel knowledgeable about the process which can help to facilitate the correct measurements.

With tools as BI and process mining, IT-systems are already collecting data, the tools can however showcase measurements missed by the systems by finding needs when looking at existing data, which is why they are presented in this step.

Level 1

Critical to Quality Tree can be used to break down customer expectations into process measurements. It can also be beneficial to have several people and brainstorm in order to get several perspectives on the matter.

Level 2

BI-Platforms can be incredibly useful when helping to define relevant measurements as it provides live data with drill-down-functionality that enables a thorough view which can be used to find what is missing and could be interesting to look at.

Level 3 and 4

Process mining can help defining measurement in several ways. The first being be defining bottlenecks and non-value adding activities within the business. The second being by comparing handmade-process maps to reality.

Do we have data?

To utilize a data driven approach access to data is critical, when solving a larger problem there are sometimes need for problem specific data that are not in standardized collection. If there is limited data and no room for extensive collection, the project can proceed with a more qualitative analysis using the process door.

Questions to answer

- What do we want to know from the data?
- Is there a data warehouse or is data available from existing systems?
- Is the data relevant and extensive?
- Is the data quality good enough?

Collect problem specific data

If there is a need to collect more data or more problem specific data, a data collection plan should be constructed to include additional metrics. To utilize the more complex data techniques, large scale continuous data collection must be in place to gather data from internal IT-systems and IoT devices.

Level 1

To assess the quality of the data, the resolution, accuracy, and precision should be evaluated. This can be done by conducting a Measurement System Analysis checking the rule of tens, conducting a Gage R&R study or an MSA drilldown.

An easy way to start is to brainstorm how to collect the data and then manually sample data into a spreadsheet solution e.g. Excel.

Level 2

Automatic collection through IT systems in the organization, data can be extracted from the systems to conduct analysis.

Level 3 and 4

Data gathering with IoT devices and sensors to collect real time data of several parameters dependent on the nature of the organization.

When there are many systems collecting data continuously and sensors and IoT devices a data warehouse can be constructed to compile data from the sources to yield more insights when analyzing.

Visualize data

When visualizing data properties to consider are distribution, variations within the process and over time, what contributes to variation and how the process is performing.

Examples of graphs and diagrams to use are

- Time series plot – to show trends and changes in the process over time
- Histograms – to show distribution of data
- Pareto charts – to show which defects contributes most to the problem
- Box or Dot plots - to show variations by process steps or categories
- Capability Analysis – to show process performance

Level 1

With small amounts of data or lack of access to specialized tools most visualizations can be performed by using Excel. However, it requires more manual tasks than specialized software.

Level 2

Statistical Software have the functions built in and therefore do not need to construct other than with simple clicks. It does however require some knowledge to understands the graphs.

When or if these visualizations are somewhat standardized to use, they could be integrated into a BI platform which can then be automatically updated. This would also enable earlier detection of some problems and thereby reduce the needed scope.

Level 3
Process Mining conformance checking can produce visualizations of bottlenecks by for example lead time or events needed per case. It can also produce more specific statistics on each process step discovered in a model or different variations of how the process is executed.

The discovered process model could provide a good foundation to build a simulation model from by showing data in each process step and transfer to a simulation software. Simulation can provide an "as-is" view of the process with statistics and a view of the flow through the process.

Level 4
Data Mining can provide an upgrade to pareto chart in the form of a Variable Importance Plot which shows the most influential variables impact on a target variable.

Analyze interesting issues or patterns from visualization
This step is performed to assure that patterns, changes, or issues in the data are significant over time and not due to random variation.

Level 1
When there is limited data and it has been analyzed to its full extent, the rest of the analysis can utilize a process door analysis where focus is on qualitative tools such as Value stream mapping and analyze value added time through workshops or interviews.

With small amounts of data or lack of access to specialized tools most analyses can be performed by using Excel. However, it requires more manual tasks than specialized software.

Level 2
Statistical Software provide significance testing through hypothesis testing, ANOVA, produce confidence intervals and can show correlations through a matrix plot

Level 3
Process Mining conformance checking can show where the process deviates from intended execution and process compliance and adherence.

The discovered process model could provide a good foundation to build a simulation model from by showing data in each process step and transfer to a simulation software. Simulation can provide an "as-is" view of the process with statistics and a view of the flow through the process.

Level 4
Data Mining algorithms can be used to find previously unknown and statistically significant patterns, correlations, and influential variables in the data. Examples of algorithms are decision trees, association, cluster analysis.

Root Cause Analysis

This step is performed to investigate causality in the statistically significant patterns found in the previous step. Depending on the nature of the project the tools can be used in different ways and they should be applied together as much as possible. The more data and specific analyses performed the closer to the root causes it is possible to get.

Level 1

Ishikawa diagram and 5 Whys should be used together with all tools identify causes and problems that might not be present in data as well as issues found in the data and should be used on all levels. There are several ways to digitize such as drawing and modelling tools.
Design of Experiments is a way to create transfer functions between inputs and outputs of the process.

Level 2

Statistical software can be used for hypothesis testing and further statistical analysis to confirm dependence between variables.

Level 3

Process mining can be used to conduct similar analysis as statistical software with a true view of every variation of the process recorded in the event log.

Level 4

Data mining and machine learning can provide parameter settings or show most influential variables of both numerical and categorical predictor variables which can prevent the need to conduct a designed experiment.

Simulations on a digital twin can be performed to conduct root cause experiments by altering settings in the digital model.

Generate solutions

More data at hand enables getting closer to causes and can provide directions for the improvements, however solutions by brainstorming or possible engineering fixes are common methods to find improvements. Only the higher-level tools can provide concrete actions to take based on the root cause analysis.

Compare solution options impact on customer needs and process requirements according to criteria set by the project team. A cost/benefit analysis should be performed to make sure solutions are economically feasible to implement.

Level 1 and 2

Brainstorming and possibilities for error proofing mechanisms are common methods to generate solutions.

Level 3

Process mining can help pinpoint automation opportunities for Robotic Process Automation.

“What if” simulation with the proposed changes built into the simulation model to test the solutions. Results of simulations should be compared to the results from the “as-is” simulation.

Level 4
Data mining and machine learning can provide parameter settings or show most influential variables of both numerical and categorical predictor variables and tune the variables accordingly.

By using a digital twin of the process more credible simulations can be performed as it is based on real time data from the actual process.

Implementation

When implementing a solution there are several aspects and risks to consider. To assess risks FMEA can be used to consider both process risks and implementation risks. Implementation plans should be set up for all sub projects of implementing each solution. A pilot project on small scale are a good way to test the solution in action before scaling up.

Follow up

It is very important to conduct the follow up to see if the changes to the process has in fact been an improvement and if they are now part of the process as well if lessons learned can be applied to other processes. Therefore, a plan for monitoring the process according to at least the measurements previously defined in the project. If there is no recorded improvement, go back to generating solutions and take another look at the root causes.

Questions to answers
- Is there a measured improvement?
- Are the measurements consistently collected?
- Have the solutions been accepted and followed?

Level 1
Measurements can be monitored through spreadsheets

Level 2
Statistical Software can be used to construct control charts to monitor the process through.

Trends and measurements of progress is where BI platforms perform best, automatic update and monitoring can be set up to enable follow up

Level 3
Process Mining can enable follow up by continuous loading of data into the process model to monitor progress and process variation. A dashboard with the key measurements can be a way to enable follow up.

Level 4
Data mining or machine learning algorithms developed in the project can keep running and make use of automated alerts of deviations.

Done

The problem was solved, it is now important to give feedback to the person who came up with the improvement proposal in order for them to see that things have changed for the better. Feedback and letting everybody be committed is an essential part to building a quality culture.

7.2 Smaller framework for Continuous Improvements

The second structure is intended for small problems or improvement opportunities that occur as agitations or suggestions raised by people in the process. It excludes some of the comprehensive steps from the larger structure as this framework is aimed to be applied to changes that are cheap, fast and is intended to build a quality culture. The shortcuts taken are possible if for example the cost in form of time spent on the investigation outweigh the cost for an incorrect change. An overview of the small structure is shown in Figure 28.

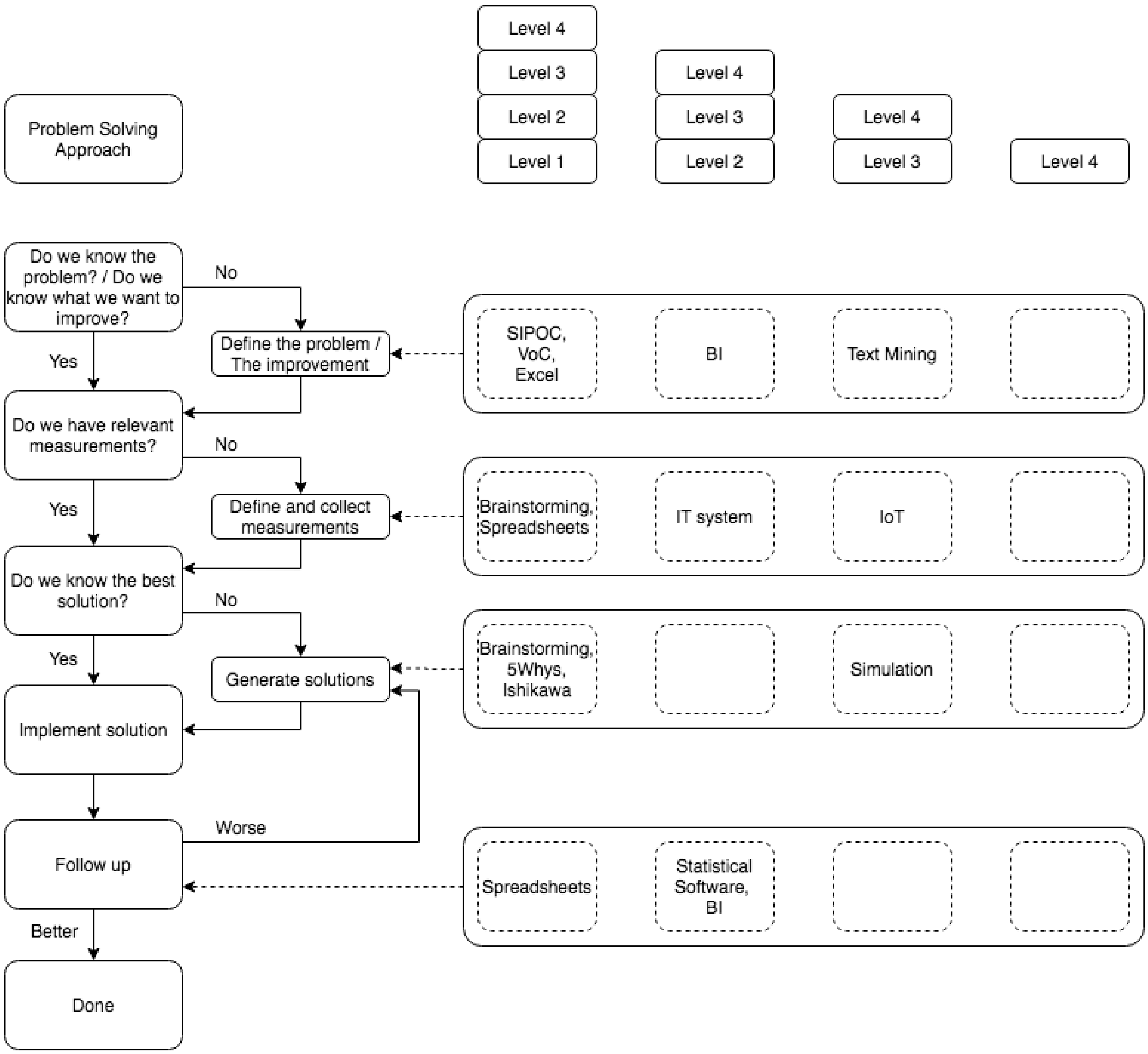

Figure 28. Smaller framework for continuous improvements

Do we know the problem? Do we know what we want to improve?

The first part of the framework is to make sure the scope of the improvement is understood and clearly defined as well as who should be in the team.

Questions to answer

- What is the aim with the improvement?
- What process is the improvement connected to?
- Is there need for a team, if so, who should be in it?

Define the problem / the improvement

This phase aims to help answer the questions from the previous phase if they failed to be satisfied. The tools that can be used to define the improvement for each maturity level are explained below.

Level 1

Observations and interviews can be used to understand the aim or need for the improvement. This can also help to understand how the process is connected. Tools like SIPOC and VOC can also be used when defining customers and understanding them. Finally, a brainstorming session or a knowledge matrix can be used to determine who should be in the team.

Level 2

Standardized report in a BI-platform can help to clearly scope and define an aim of the improvement as well as indicate current trends.

Level 3 and 4

Text mining on customer reviews can be used to improve understand the voice of the customer better.

Do we have relevant measurements?

This is one of the most important parts in the framework. The only way of knowing whether a change was an improvement or failure is to measure the output before and after. It is therefore important to have relevant measurements for the specific improvement project.

Questions to answer

- Do we have the correct measurements for the project?
- What measurement indicate an improvement??

Define and collect measurements

If there are no relevant measurements for the process it is time to define and start collecting them.

Level 1

An easy way to start is to brainstorm which measurements that could be relevant and begin measuring them by using spreadsheets e.g. Microsoft Excel.

Level 2

Automatic collection through IT systems in the organization, data can be extracted from the systems to indicate progress.

Level 3 and 4
IoT-sensors can be installed around a business. These sensors are able to collect enormous amount of live data which can be used to get an even more precise description of reality.

Do we know the best solution?
If there are several solutions it is important to reflect if the proposed solution is the best solution and not just the first or the easiest.

Questions to answer
- Do we have a solution?
- Is this the best possible solution for the suggested improvement?

Generate solutions
A workshop is a great way of solutions generation or deciding on the best solution. As this framework is intended for small alterations the impacts of changing the process should be limited. Therefore, a trial and error approach can be suitable.

Level 1 and 2
Brainstorming is an effective way of generating solutions. If there are several solutions 5 whys in combination with an Ishikawa diagram can be powerful when deciding the best solution.

Level 3 and 4
If a simulation model can be constructed quickly, it could be worthwhile to test implementation to determine the best solution. However, this is more likely used in the larger framework for bigger problems.

Implement solutions
The step when the selected solution is implemented. This can be performed in many different ways depending on what the implementation will be. The important thing to have is a leader who has the power to implement the change.

Questions to answer
- How do we implement the change?

Follow up
This is one of the most important steps in the improvement process and is oftentimes the part that gets neglected. It is time to compare the measurements from phase 2 “relevant measurements” and see if the change resulted in an improved or a worsened performance. If it is an improvement the project can be deemed a success, otherwise go back to generating solutions and reiterate.

Questions to answer
- How has the measurement changed?

Level 1
An easy way of following up is to compare the measurements in the spreadsheet, e.g. Microsoft Excel.

Level 2, 3 and 4
The measurements can also be compared in a statistical software to assure that the change is significant and not by chance.

An incredibly powerful tool for follow up is BI platforms. It gives instant feedback as the software can be fed with live data. The changes effect can be seen directly.

Done
The improvement is deemed a success, it is now important to give feedback to the person who came up with the improvement proposal in order for them to see that things have changed for the better. Feedback and letting everybody be committed is an essential part to building a quality culture.

8 Discussion and Conclusions

This part consists of a discussion of the results, methods used, and conclusions drawn in regard to the research questions of the study.

8.1 Discussion of results

As the use cases and unique applications of the tools can differ greatly depending on factors such as collected data or type of process, there are need for interpretation and to include a variety of competencies to utilize the framework in an effective manner. A significant aspect to consider is regarding the variation in specific use cases and usefulness of different tools dependent on the industry a company operates within. The framework is however based on a well-established improvement structure and could be mapped against the existing improvement processes that has been used in both manufacturing and service, presented in Figure 4 with their respective phases and flexibilities in applying tools. The framework was revised and developed with input from five different organizations, with different industries represented, this provided an empirical range of tools from different maturity levels that validate an ongoing maturing process to integrate more tools into existing structures. The benefits reaped from specific tools can however fluctuate vastly and there is “no size fits all” when choosing what tool to implement.

The digital maturity model presented in this thesis could be viewed as a measure of maturity in data analysis and seen as a sub-model of digital maturity. Digital transformation extends to involve transforming business models and organizational aspects which the digital maturity model presented does not explicitly incorporate. It does however attempt to incorporate the connection to process maturity which has been validated to play a role in better use of digital tools.

The need for updated team structures with closer collaboration between departments as well as continuous training in new topics was highlighted in both theory and empirical findings. One of the delimitations of the study was the back-end structuring of data and integration of information systems. These are significant parts in using many of the tools and often the most time consuming and complex part. Currently various systems within organizations are being cleaned and integrated to create data warehouses that can be used by the various digital tools. The analytical software providers in turn are developing their platforms so that they contain more and more functions and pre-existing algorithms. It is therefore authors belief that it will be more important to know the purpose for use and logic of applying the tools to be able to take advantage and make better decisions.

8.2 Discussion of Method

A large part of theoretical base studied on application of the tools was in connection with presented improvement processes and in particular Six Sigma. This could have been broadened to several other applications to include even more use cases, it was however deemed that the resulting framework would still require much interpretation to use. The qualitative methods used were preferable to gain specific insight in use cases and depth in discussions with respondents. Due to the variety of the organizations, ranging from local to industry leading global companies, the empirical findings provided a holistic view of the subject. The interviews were forced to be changed from physical to digital interviews on short notice due to COVID-19. Questions were designed to be open and enable a discussion with the possibility for follow up questions. The digital format might have benefitted from more structured interviews guide than semi structured as follow up questions and reading of body language was somewhat impaired due to delay in some interviews.

Both authors of the study have a background in quality management and were unfamiliar with many of the topics presented. Personal bias towards simplicity in quality tools might be present when establishing levels for the various tools that accompanies the improvement structures in the result. Concerns raised in literature did however act as indications and to further cope with this, the suggested levels were presented at a discussion seminar with the aim of getting input on the final result

8.3 Conclusions

The aim of this study has been to analyze and present how digital tools can be incorporated into structured improvement work while taking a business digital maturity into consideration. This was performed by answering the three research questions.

Research Question 1

What steps of the data analysis and improvement process lack or have limited support by the digital tools of today?

The first research question laid the foundation by mapping where digital tools can support improvement work but more importantly where it lacks. The question is complicated to answer as digital tools can be incorporated into improvement structures in several different ways. It depends on the type of business and its specific needs, the business digital maturity as well as what type of structure is used. Digital tools can theoretically be implemented in all steps in an improvement process, however most times it is not a viable option. Some steps require high level complex tools as data mining and machine learning which most business cannot reap the benefits of and therefore are better suited to use simpler tools. Root cause analysis is a step where it is possible to utilize data mining to narrow down parameters and closing in on the root cause but still need to complement with analogue tools such as Ishikawa and 5whys in order to determine the root cause. It is theoretically correct to state that digital tools can act as support to all steps in an improvement structure, however practically most business cannot utilize the higher-level tools.

Research Question 2

How can today’s digital tools be used in a structured analysis and improvement process?

Research Question 3

How can a structured analysis and improvement process be adjusted to match a business’ digital maturity?

The second and third research questions builds on the conclusions from the first research question. As previously mentioned, digital tools have vastly different complexity and therefore needs to be matched to specific digital maturity levels for a framework to be useful. This was answered in the form of the final frameworks presented in *Section 7* which presents two structured methodologies for improvement work and problem solving. The frameworks also take digital maturity into account by suggesting maturity levels for the digital tools to be incorporated. The result is one possible way of incorporating digital tools into structured improvement work, it is possible to use other improvement structures with the same basic thought of integrating digital tools in different maturity levels.

8.4 Future research

The digital and process maturity model should be further investigated and maybe complemented by a diagnostic tool similar to the accompanying tool for process maturity for the digital maturity side of the model to display a clear roadmap to follow to reach higher maturity levels.

The road map through the improvement structure has a broad theoretical foundation, however the rankings of the specific tools could use a more rigorous and theoretically based ranking through further empirical studies with additional businesses. To put in new tools at various levels, clear assessment criteria would be useful. The authors have used categories such as difficulty, generalizability in use cases, depth of analysis and connection to process maturity, which could all be suitable categories to start from.

As all studies have limited time and resources and therefore delimits certain aspects, many digital tools existing on the market have not been included. During the study there has been a tradeoff between gaining an understanding of presented tools and broadening the scope of tools to include. More tools could be included and a method for assessing and implementing new tools are probably beneficial for many organizations.

9 References

Aggarwal, C. C., 2015. *Data Mining -*. Yorktown Heights, New York: Springer.

Alexanderson, K., 2016. *Internetstiftelsen.* [Online] Available at: https://internetstiftelsen.se/docs/Kallkritik-pa-Internet.pdf [Accessed 1 Februari 2020].

Ali, Z. & Bhaskar, . S. B., 2016. Basic statistical tools in research and data analysis. *Indian Journal of Anaesthesia*, September, Volume 60, p. 662–669.

ALLEA - All European Academies, 2017. *The European Code of Conduct for Research Intregrity*, Berlin: ALLEA - All European Academies.

Alp, U. & Emre, C., 2018. *Industry 4.0: Managing The Digital Transformation.* 1:st edition ed. Switzerland: Springer.

Alvarado, J., Aguirre, S. & Parra, C., 2012. Combination of Process Mining and Simulation Techniques for Business Process Redesign: A Methodological Approach. *International Symposium on Data-Driven Process Discovery and Analysis,* 162(1), pp. 24-43.

Andersen, B. & Natland Fagerhaug, T., 2006. *Root Cause Analysis: Simplified Tools and Techniques*, Milwaukee: ASQ Quality Press.

Anthony, J. et al., 2019. An evaluation into the limitations and emerging trends of Six Sigma: An Empirical Study. *The TQM Journal*, 31(2), pp. 205-221.

Azeroual, O. & Theel, H., 2019. The Effects of Using Business Intelligence Systems on an Excellence Management and Decision-Making Process by Start-Up Companies: A Case Study. *International Journal of Management Science and Business Administration*, March, 4(3), pp. 30-40.

Barone, S. & Franco, E. L., 2012. *Statistical and Managerial Techniques for Six Sigma Methodology.* 1:st edition ed. West Sussex: John Wiley and Sons, Ltd.

Bell, J., 2016. *Introduktion till forskningsmetodik.* 5:e ed. Lund: Studentlitteratur AB.

Bergman, B. & Klefsjö, B., 2010. *Quality from Customer Needs to Customer Satisfaction.* 3 ed. Lund: Studentlitteratur.

Björklund, M. & Paulsson, U., 2012. *Seminarieboken - Att skriva, presentera och opponera.* 2:a ed. Lund: Studentlitteratur AB.

Boyd, D. & Crawford, K., 2011. Six Provocations for Big Data. *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*, 21 September, pp. 1-3.

Broman, K. W. & Woo, K. H., 2018. Data Organization in Spreadsheets. *The American Statistician*, 24 April, 72(1), pp. 2-12.

Brook, Q., 2017. *Lean Six Sigma and Minitab: The Complete Toolbox Guide for Business Improvement.* 5:e ed. s.l.:Opex Resources Ltd..

Calude, C. S. & Longo, G., 2016. The Deluge of Spurious Correlations in Big Data. *Foundations of Science*, 7 March, 59(7), pp. 595-612.

Chanias, S. & Hess, T., 2016. *How digital are we? Maturity models for the assessment of a company's status in the digital transformation,* Munchen: LMU Munich.

Chapman, P. et al., 2000. *CRISP-DM 1.0 - Step-by-step data mining guide,* s.l.: SPSS Inc..

Chu , X., F. Ilyas, I., Krishnan, S. & Wang, J., 2016. *Data Cleaning: Overview and Emerging Challenges.* New York, NY, USA, Association for Computing Machinery, p. 2201–2206.

Claes, J. & Poels, G., 2012. *Process Mining and the ProM Framework: An Exploratory Survey.* Ghent, Business Information Processing.

Cronemyr, P., 2007. DMAIC and DMADV - differences, similarities and synergies. *International Journal of Six Sigma and Competitive Advantage*, 3(3), pp. 193-209.

Cronemyr, P. & Danielsson, M., 2013. Process Management 1-2-3 - a maturity model and diagnostics tool. *Total Quality Management & Business Excellence,* 24(7-8), pp. 933-944.

DAMA UK Working Group, 2013. *THE SIX PRIMARY DIMENSIONS FOR DATA QUALITY ASSESSMENT Defining Data Quality Dimensions,* United Kingdom: DAMA UK Working Group.

Dasu, T. & Johnson, T., 2003. *Exploratory Data Mining and Data Cleaning.* 1:st ed. New Jersey: John Wiley & Sons Inc.

De Mauro, A., Grimaldi, M. & Greco, M., 2014. *What is Big Data? A Consensual Definition and a Review of Key Research Topics.* s.l., s.n.

Duarte, J., 2017. Data Disruption. *Quality Progress,* pp. 20-24.

Eckerson, W. W., 2007. *PREDICTIVE ANALYTICS Extending the Value of Your Data Warehousing Investment,* Chatsworth CA: TDWI.

EMC education Services, 2015. *Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data.* 1:st edition ed. Indianapolis: John Wiley & Sons, Inc..

Evans, J., 2012. Business Analytics: The Next Frontier for Decision Sciences. *Decision Line*, March, 43(2), pp. 4-6.

Fahey, W., Jeffers, P. & Carroll, P., 2020. A business analytics approach to augment six sigma problem solving: A biopharmceutical manufacturing case study. *Computers in Industry,* Volume 116.

Fahmy, A. F., Yousef, A. H. & Mohamed, H. K., 2017. *A Data Mining Experimentation Framework to Improve Six Sigma Projects.* Cairo, IEEE.

Gallupe, R. B. & Cooper, W. H., 1993. Brainstorming Electronically. *MIT Sloan Management Review*, 35(1).

Gartner, Inc. , 2020. *Gartner Glossary.* [Online] Available at: https://www.gartner.com/en/information-technology/glossary/diagnostic-analytics [Accessed 28 March 2020].

George, M., Rowlands, D. & Kastle, B., 2004. *What is LEAN SIX SIGMA?.* 1:st edition ed. New York: McGraw-Hill.

Geyer-Klingeberg, J., Nakladal, J., Baldauf, F. & Veit, F., 2018. *Process Mining and Robotic Process Automation: A Perfect Match.* Sydney, s.n.

Ghosh, S. & Maiti, J., 2014. Data mining driven DMAIC framework for improving foundry quality - a case study. *Production Planning & Control,* 25(6), pp. 478-493.

Godfrey, A. B., 2008. Eye on data quality. *Six Sigma Forum Magazine,* Volume 8(1), pp. 5-6.

Gray, P. & Negash, S., 2008. Business Intelligence. In: F. Burstein & C. W. Holsapple, eds. *Handbook on Decision Support Systems 2.* Berlin: Springer, pp. 175-195.

Gupta, S., Modgil, S. & Gunasekaran, A., 2020. Big data in lean six sigma: a review and further research directions. *International Journal of Production Research,* 58(3), pp. 947-969.

Hamra , G. B. & Schinasi, L. H., 2017. A Time Series Analysis of Associations between Daily Temperature and Crime Events in Philadelphia, Pennsylvania. *Journal of Urban Health volume*, 07 07, pp. 894-899.

Jukić, N. & Nestorov, S., 2006. Comprehensive data warehouse exploration with qualified association-rule mining. *Decision Support Systems*, November, pp. 859-878.

Jørgensen, P. S. & Rienecker, L., 2014. *Att skriva en bra uppsats.* 1:st edition ed. Stockholm: Liber.

Kane, G. C., 2017. *Digital Maturity, Not Digital Transformation.* [Online] Available at: https://sloanreview.mit.edu/article/digital-maturity-not-digital-transformation/ [Accessed 9 March 2020].

Koncz , A. & Pokorádi , L., 2018. *8D Usage in Automotive Industry.* Budapest, IEEE.

Kritzinger, W. et al., 2018. Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine,* 6 September, pp. 1016-1022.

Laux, C., Li, N., Seliger, C. & Springer, J., 2017. Impacting Big Data analytics in higher education through Six Sigma techniques. 66 (5), pp. 662-679.

Lee, J., Kao, H.-A. & Yang, S., 2014. Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment. *Procedia CIRP*, pp. 1-3.

Liesener, T., 2013. *Kaizen Factory.* [Online] Available at: https://kaizen-factory.com/2013/09/11/pdca-a3-dmaic-8dpsp-what-are-the-differences/
[Accessed 26 Februari 2020].

Magnusson, K., Kroslid, D. & Bergman, B., 2003. *Six Sigma The Pragmatic Approach.* 2:nd edition ed. Sweden: Studentlitteratur.

Mansharamani, R., 1997. An overview of discrete event simulation methodologies and implementation. *Sadhana*, October, 22(1), pp. 611-627.

Marr, B., 2018. *How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read,* s.l.: Forbes.

Matthews, D. D., 2011. *The A3 Workbook: Unlock Your Problem-Solving Mind.* 1:st edition ed. New York: Auerbach Publications.

Melan, E. H., 1993. *Process management: methods for improving products and service.* New York: McGraw Hill, Inc..

Montgomery, D. C., 2013. *Statistical Quality Control: A modern Introduction.* Seventh Edition ed. Singapore: John Wiley & Sons.

Pajo, B., 2017. *Introduction to research methods: a hands-on approach.* Los Angeles: SAGE Publications.

Propia AB, 2019. *Propia.* [Online] Available at: https://www.propia.se/Examensarbete.html [Accessed 22 Januari 2020].

Propia, 2020. [Interview] (15 March 2020).

Riesenberger, C. A. & Sousa , S. D., 2010. The 8D Methodology: An Effective Way to Reduce Recurrence of Customer Complaints?. *Proceedings of the World Congress on Engineering*, 2 June.

Saunders, M., Thornhill, A. & Lewis, P., 2009. *Research Methods for Business Students.* 5:th edition ed. Edinburgh: Pearson Education Limited.

Schober, P., Boer, C. & Schwarte, L. A., 2018. Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia & Analgesia*, 05, pp. 1763-1764.

Schäfer, F., Zeiselmair, C., Becker, J. & Otten, H., 2018. *Synthesizing CRISP-DM and Quality Management: A Data Mining Approach for Production Processes.* Marrakech, IEEE.

Sehlin, D., Truedsson, M. & Cronemyr, P., 2019. A conceptual cooperative model designed for processes, digitalisation and innovation. *International Journal of Quality and Service Sciences,* 11(4), pp. 504-522.

Selvam, V. & Babu, R., 2015. An Overview of Machine Learning and its Applications. *International Journal of Electrical Sciences & Engineering (IJESE)*, January, 1(1), pp. 22-24.

Sharda, R., Delen, D. & Turban, E., 2014. *Business Intelligence A managerial Perspective on Analytics.* 3:d edition ed. Essex: Pearson Education Limited.

Simon, P., 2017. *Analytics: The Agile Way,* s.l.: John Wiley & Sons.

Steils, N. & Hanine, S., 2019. Value-Added Crowdsourcing: Digital Catalysts for Creative Contests. In: *Managing Diversity, Innovation, and Infrastructure in Digital Business.* s.l.:IGI Global, pp. 160-179.

Su, C.-T., 2019. Take a Big BYTE. *Quality Progress*, October, pp. 22-29.

Sun, J.-T. & Cai, Y., 2009. Text Mining. In: L. LIU, ed. *Encyclopedia of Database Systems.* Boston, MA: Springer US, pp. 3061-3065.

Swedish Research Council, 2017. *Good research practice,* Stockholm: Swedish Research Council.

Swedish Research Council, 2019. *Ethics in research.* [Online] Available at: https://www.vr.se/english/mandates/ethics/ethics-in-research.html [Accessed 14 February 2020].

Tang, N., 2014. Big Data Cleaning. *Web Technologies and Applications,* Issue 16, pp. 29-40.

Teichert, R., 2019. Digital Transformation Maturity: A Systematic Review of Literature. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis,* 67(6), pp. 1673-1687.

Vaisman, A. & Zimányi, E., 2014. *Data Warehouse Systems: Design and Implementation.* 1:st edition ed. Berlin: Springer.

Van den Broeck, J., Argeseanu Cunningham, S., Eeckels, R. & Herbst, K., 2006. Data Cleaning: Detecting, Diagnosing, and Editing Data Abnormalities. *PLOS Medicine*, 6 September, p. 966.

van der Aalst, W., 2016. *Process Mining - Data Science in Action.* 2:a ed. Berlin: Springer-Verlag Berlin Heidelberg.

van der Aalst, W. M. P., Bichler, M. & Heinzl, A., 2018. Robotic Process Automation. *Business & Information Systems Engineering – Springer Journals*, 60(4), pp. 269-272.

Vigen, T., 2015. *Spurious Correlations.* 1:st edition ed. 2015: Hachette Books.

Xiong, C., Shapiro, J., Hullman, J. & Franconeri, S., 2019. Illusion of Causality in Visualized Data. 01 08, pp. 1-2.

Zwetsloot, I. M., Kuiper, A., Akkerhuis, T. S. & de Koning, H., 2018. Lean Six Sigma meets data science: Integrating two approaches based on three case studies. *Quality Engineering*, 30(3), pp. 419-431.

Appendix A

Keywords used in the literature study are presented in the list below, the terms were searched by themselves as well as in combination with each other.

- Six Sigma
- DMAIC
- Improvement Process
- Root Cause Analysis
- Process Management
- Process Maturity
- Big Data
- Analytics
- Business Analytics
- Business Intelligence
- Data Mining
- Machine Learning
- CRISP-DM
- Data Quality
- Process Mining
- Robotic Process Automation
- Digital tools
- Digital
- Simulation
- Digital Transformation
- Digital Maturity
- Digitalization

Appendix B

Interview Template divided by the three main areas, beyond these questions, follow-up questions differed between respondents dependent of the answers received.

Improvement Processes

How is improvement work generally conducted?

Do you use any particular improvement methodology, if so which one?

Is the improvement work data-driven, if so how?

What are the main obstacles you experience when driving improvement work?

(Digital) Tools

What type of tools do you utilize in your improvement work?

How do you utilize your current digital tools?

Do you see possibilities to incorporate more digital tools in your company, if so which and why?

What would you say is missing for you to work with more digital tools?

Process/Digital Maturity

How do you work with processes?

How difficult is the data used for improvement work accessed?

How digitally mature would you consider your company to be?