

St. Petersburg University
Graduate School of Management

Master in Management Program

INNOVATION TRIGGERED BY BIG DATA ANALYSIS OF MANUFACTURING
MACHINE DATA

Master's Thesis by the 2nd year student — Sebastian Grossalber, BSc.

Research advisor: Associate Professor, Dmitry Vyacheslavovich Kudryavtsev

St. Petersburg

2018

Declaration of Authorship

I hereby declare that I have written this Master Thesis myself, independently and without the aid of unfair or unauthorized resources. Whenever content has been taken directly or indirectly from other sources, this has been indicated and the source referenced. This Master Thesis has not been previously presented as an examination paper in this or any other form in Russia or abroad. This Master Thesis is identical with the thesis assessed by the examiner.


Saint Petersburg, May 23, 2018

АННОТАЦИЯ

Автор	Себастьян Гроссальбер
Название ВКР	Инновации, основанные на анализе больших данных с производственных машин
Образовательная программа	Программа «Менеджмент»
Направление подготовки	
Год	2018
Научный руководитель	Кудрявцев Дмитрий Вячеславович
Описание цели, задач и основных результатов	Большой анализ данных позволяет обнаруживать новые знания из неструктурированных данных. Эта практика уже сейчас относительно распространена и даже применяется к цепочке поставок продукции и производству в целом для снижения воздействия узких мест. Тем не менее, также через производственные машины производят значительный объем данных, эти данные во многих случаях либо не хранятся, либо не используются для запуска инноваций, основываясь на выводах данных производственной машины. Этот тезис направлен на оценку возможностей и использования производственных данных для инноваций. Используя опрос среди производственных компаний, изучаются предпосылки для инноваций, основанных на машинных данных. После этого многократное исследование исследует явления и проблемы в большей степени.
Ключевые слова	Инновации, производство, машинные данные, большие данные, оптимизация, IoT, Интернет вещей, промышленность 4.0, производство

ABSTRACT

Master Student's Name	Sebastian Grossalber
Master Thesis Title	Innovation triggered by Big Data analysis of manufacturing machine data
Educational Program	Master in Management - MIM
Main field of study	

Year	2018
Academic Advisor's Name	Dmitry Vyacheslavovich Kudryavtsev
Description of the goal, tasks and main results	<p>Big Data analysis enables to discover new knowledge out of unstructured data. This practice is already now relatively common and is even applied to production supply chain and production in general to decrease the effects of bottlenecks. However, also through production machines produce a considerable amount of data, this data is in many cases either not stored or not used for triggering innovation, based on the findings out of the production machine data. This thesis aims to evaluate the opportunities and usage of production machine data for innovation. Using a survey amongst manufacturing companies the preconditions for machine data-based innovation are explored. Following that a multiple case study investigates the phenomena and challenges in greater depth.</p>
Keywords	Innovation, manufacturing, machine data, Big Data, optimization, IoT, Internet of things, Industry 4.0, production

Abstract

Big Data analysis enables to discover new knowledge out of unstructured data. This practice is already now relatively common and is even applied to production supply chain and production in general to decrease the effects of bottlenecks. However, also through production machines produce a considerable amount of data, this data is in many cases either not stored or not used for triggering innovation, based on the findings out of the production machine data. This thesis aims to evaluate the opportunities and usage of production machine data for innovation. Using a survey amongst manufacturing companies the preconditions for machine data-based innovation are explored. Following that a multiple case study investigates the phenomena and challenges in greater depth.

Acknowledgment

I would first like to thank my parents and my whole family for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Furthermore, I must express my very profound gratitude to my thesis advisor Dmitry Vyacheslavovich Kudryavtsev of the Graduate School of Management at the St. Petersburg State University whose advice and feedback helped me tremendously in my research.

- Sebastian Großalber

ABSTRACT	5
ACKNOWLEDGMENT	6
LIST OF FIGURES	9
CHAPTER 1: INTRODUCTION	10
BIG DATA	10
BIG DATA & PRODUCTION	11
INTERNET OF THINGS	12
PRODUCTION & PREDICTIVE MAINTENANCE	13
INNOVATION	15
DIFFERENTIATION OF OPTIMIZATION AND (DISRUPTIVE-) INNOVATION	16
RESEARCH GAP	17
STRUCTURE	19
CHAPTER 2: METHODOLOGY	20
TYPE OF RESEARCH	20
RESEARCH STRATEGY	20
SURVEY QUESTIONS	21
INTERVIEW	27
CHAPTER 3: SURVEY RESULTS	28
FINDINGS	29
SAMPLE CHARACTERISTICS	35
IMPACT OF SURVEY RESULTS ON THE RESEARCH QUESTIONS	37
CHAPTER 4: MULTIPLE CASE STUDY	38
COMPANY 1	38
COMPANY 2	40
COMPANY 3	41
COMPARISON	43
CHAPTER 5: CONCLUSION	49
CHAPTER 6: LIMITATIONS	51
CHAPTER 7: MANAGEMENT IMPLICATIONS	52
BIBLIOGRAPHY	56
APPENDIX	59

SURVEY QUESTIONS	59
INTERVIEW WITH INTERVIEW PARTNER A FROM COMPANY 1 (INTERVIEW C1A)	65
INTERVIEW WITH INTERVIEW PARTNER B FROM COMPANY 1 (INTERVIEW C1B)	68
INTERVIEW WITH INTERVIEW PARTNER A FROM COMPANY 2 (INTERVIEW C2)	74
INTERVIEW WITH INTERVIEW PARTNER A AND B FROM COMPANY 3 (INTERVIEW C3)	78

List of figures

FIGURE 1 COMPARISON OF WHICH RESEARCH HAS BEEN PERFORMED REGARDING OPTIMIZATION AND INNOVATION	17
FIGURE 2 INTERNAL DATA BEING USED FOR THE CREATION OF NEW PRODUCTS	29
FIGURE 3 IS CUSTOMIZATION PERFORMED TO GENERATE DATA?	29
FIGURE 4 GOAL OF DATA GENERATION	30
FIGURE 5 REGULAR STORAGE	31
FIGURE 6 OPEN OR CLOSED ANALYSIS OF DATA	31
FIGURE 7 ANALYSIS TECHNIQUES BEING USED	32
FIGURE 8 ACCESS TO DATA	32
FIGURE 9 ARE EXPERIMENTS PERFORMED?	33
FIGURE 10 TYPES OF EXPERIMENTS PERFORMED.	33
FIGURE 11 EXPERIMENT ENVIRONMENT	34
FIGURE 12 HOW DOES DATA HELP REGARDING PRODUCTS?	34
FIGURE 13 HOW MUCH DOES DATA HELP REGARDING CREATING NEW PRODUCTS?	35
FIGURE 15 SHARE OF COMPANIES BY EMPLOYEE SIZE CLASSES	35
FIGURE 16 NUMBER OF EMPLOYEES	36
FIGURE 17 INDUSTRY	36
FIGURE 18 LOCATION OF HEADQUARTERS OR MAJOR MARKET	37

Chapter 1: Introduction

To survive over time, it is necessary for a company to innovate continually and further develop their existing portfolio. Moreover, modern production machines are generating vast amounts of data points every minute of the production process. Combining the two, machine data and innovation, would enable companies to take advantage of the data masses produced inside their organization every day to be more innovative than their competition and thereby to generate a sustainable strategic advantage.

In this thesis, I try to evaluate if and how companies are leveraging machine data analysis to boost their innovation. The focus lies on the multiple case study that shows how companies that already use machine data for their innovation processes (or are trying to do). These best practices give managers a first insight into the possibilities and issues that are linked to machine data analysis triggered innovation.

Based on the multi-field nature of the topic, several scholars from different research fields are used to evaluate the influence of production data based Big Data analysis on innovation.

Big Data

Nowadays the amounts of data collected in companies exceed their capabilities of what can analyze by far. Big Data analysis tools can help to make sense of all this data and even facilitate the analysis of even higher amounts of data. (McAfee, Brynjolfsson, & Davenport, 2012)

Studies show that companies are starting to focus more and more on Big Data and BI&A (business intelligence & analytics) (Chen, Chiang, & Storey, 2012). Therefore, it is not surprising at all, that Big Data is one of the major trends (as well as a buzzword) for companies as well as for research.

To analyze the possibilities of this trend, first the term has to be defined:

„Big Data is a term for massive data sets having large, more varied and complex structure with the difficulties of storing, analyzing and visualizing for further processes or results. “
(Sagiroglu & Sinanc, 2013)

Big Data allows to perform analysis that cannot be done on a small scale, so it is more than just the same analysis that is already done on data for years. Combining the different data points allows extracting utterly new information from the dataset. This information can generate new forms of value and to change how markets, institutions, and governments work. (Cukier & Mayer-Schoenberger, 2013)

Established companies usually do have many data already, or it is easy for them to start collecting new data, so the primary challenge for these companies is not the gathering of Big Data, but the analysis and what to do with it. This is where Big Data Analytics gets into the game. Due to statistical and data mining techniques (e.g., text analytics, web analytics, network analytics and mobile analytics) it is possible to generate meaningful results out of tremendous amounts of data points. (Chen et al., 2012) These results may vary from analyzing a person's behavior in the past up to predict future decisions.

Business can utilize these results for various applications: from customer-related applications where the organization targets new customers or services existing ones better, to organization related topics like smart supply, smart production or all together: a smart factory.

Big Data & Production

One of the core elements of Big Data & Production (so-called Industry 4.0) is cyber-physical systems (CPS). In a nutshell, CPS is a connection between virtual and physical production elements. The employees, the production system, the products and in the end even customers are connected to each other. Those elements are the basis for intelligent and self-reconfiguring production systems. The so-called smart factory uses former workers as operators, which only monitor the manufacturing process and through the automated transfer and analysis of data, the production itself is not only steered but also continuously

and automatically improved over time. Using data for optimizing the production is one of the most significant value creation opportunities for Big Data application. (Dombrowski, Riechel, & Evers, 2014; Karre, Hammer, Kleindienst, & Ramsauer, 2017)

Industry 4.0 will, due to the connection of cyber-physical systems, enable real-time monitoring of procedures. Therefore, fewer operations will be extensively planned, and more and more tasks will be ad-hoc and trigger by individual events. (Spath et al., 2013)

Already, nowadays an unusually high amount of effort is needed for the organization and scheduling of employees. This might increase even more, due to the higher demand for flexible production and more significant fluctuations due to higher network dependencies. However, studies have shown, that organizations, which rely on self-organization of their employees, reduce coordination effort to a minimum. Consequently, a higher amount of self-organization can be expected due to Industry 4.0. (Spath et al., 2013)

Internet of Things

In 2008 there were more devices on the Internet than human beings on earth. In 2020 the figure is estimated to rise to more than 50 billion. These immense amounts of tools lead to the development of the Internet of Things (IoT). (Swan, 2012) It is based on the EPC Network (Electronic Product Code), an open technology infrastructure which enables that IT systems can identify other objects. The technical ground for this development is RFID (Radio Frequency Identification). (Sundmaeker, Guillemin, Friess, & Woelfflé, 2010)

An essential requirement for the advancement of the Internet of Things was the massive improvement in the field of hardware development which made them practical relevant for ordinary usage. The main factors have been a decline in size, lower price and better energy utilization. (Payne & Macdonald, 2006)

In principle, the Internet of Things is a network of computers and natural objects connected to each other. The IoT is not designated as a replacement of the Internet but as an addition to it. It is often described as „a shift from computer information processing to computer

sensing." (Sundmaeker et al., 2010) These networks can be built between objects like home appliances, buildings or cars, but also between the human body and the Internet. In this way, the world around but also inside the human can be researched through monitoring, as it was never possible before. (Swan, 2012)

New standards like the Narrowband Internet of Things (NB-IoT) that targets ultra-low-end IoT applications will increase their coverage as well as significantly improve the battery life. (Y.-P. E. Wang et al., 2017) Not only reduces this the costs but even more importantly allows for much more extensive applications (e.g., long-term monitoring when the sensor is not accessible after it is built in)

Production & Predictive Maintenance

Manufacturing is one of the most critical industries of many countries in Europe. The share of the economic sector in gross value added in Germany in 2016 of manufacturing companies was 25,7%. (Statista.com, 2017)

Multiple tradeoffs usually characterize production. These trade-offs include Flexibility versus productivity, standardization versus customer-oriented production, high utilization versus throughput time reduction, and so on. (Skinner, 1969) Another tradeoff is optimization versus innovation, and traditionally the data gathered by production machines is (if it is even used – at least the vast amounts of data that are available nowadays) only used for optimization (product and process optimization). However, the use of production data used for innovation is sporadic.

Any failure of production equipment reduces the productivity of a company, causes costs (e.g., if other production lines or even full plants cannot produce any more) and jeopardizes delivery dates. Nevertheless, in many cases, repair is performed only in the event of a malfunction. To reduce the risk and costs of unexpected production machine malfunction predictive maintenance (preventive maintenance based on data) can be used.

Advances in cloud computing reshape manufacturing into dynamically scalable, demand-oriented, and very distributed cost-efficient business model. Nevertheless, it also poses troubles such as dependability, availability, adaptability, machine- and information security. However, already today newly emerging Big Data and cloud analysis techniques like basing predictive maintenance on mobile agent technology (Mobile agent: originating from the field of distributed artificial intelligence, allows for operating separate analysis tasks in parallel (Cucurull et al., 2009) enable more accurate data collection, analysis and higher accuracy of malfunction prediction. (J. Wang, Zhang, Duan, & Gao, 2017)

To create a predictive system, first, a knowledge base must be acquired based on data from the past. This is done to prepare forecasts for further development based on this. A conventional approach used in this training is neural networks. Neuronal networks calculate an input quantity to an output using a variable number of connection nodes. The weighting of the intermediate nodes can be trained via a backward calculation. Furthermore, by using a brute force test method, the number of connection nodes can also be determined (by using the best result from a tested interval). All information now calculated is managed in a Knowledge Base. The determined state of the machine and the comparison of the current machine data with an anticipated normal state, in turn, make it possible to make a prediction. Production optimization can have either a quantified or a qualitative objective. While the goal of the quantified approach is to maximize availability at given maintenance costs, the qualitative approach focuses on reducing maintenance costs while maintaining availability. (Matzkovits, Saumer, Steinbach, Zielke, & Seitz, 2017)

Testing predictive maintenance techniques against regular time based scheduled maintenances, shows a significant advantage in process safety and efficiency. Moreover, it will save replacement time and costs, and most crucial system availability of the production line or plant. (Hashemian & Bean, 2011)

As the approach to creating a predictive system is well researched and tested it is not a surprise that there are many business applications using data to forecast malfunctions and aligning preventive maintenance accordingly. However, depending on specifications of the planned implementation of predictive maintenance the improvements and utility can vary

greatly. Therefore, before implementing the possible cost and risk reduction has to be evaluated. (Gebler, Hicks, Harrison, Barker, & Stirling, 2016)

Innovation

Innovation has been defined and redefined over the years:

Drucker's definition, *"Innovation is the specific tool of entrepreneurs, the means by which they exploit change as an opportunity for a different business or service."* (Drucker, 2014) focuses on using change in a different setting and thereby creating an opportunity. This worked somewhat as a basis for the understanding of what innovation is.

Following Drucker, Utterbach introduces the factor of success: *"Innovation is the successful exploitation of new ideas"* (Utterback, 1994)

After that Albury's definition defines the levers of innovation more plainly: *"Successful innovation is the creation and implementation of new processes, products, services and methods of delivery which result in significant improvements in outcomes, efficiency, effectiveness or quality"* (Albury, 2005)

For Moore & Hartley the structural improvement was the crucial part of innovation: *"Innovation is the successful development, implementation and use of new or structurally improved products, processes, services or organizational forms"* (Moore & Hartley, 2008)

Using the term added value in his definition *"Innovation is something new being realized with (hopefully) added value"* (Jacobs & Snijders, 2008) Jacobs connects a buzzword that is used in so many business discussions to innovation.

However, even though there is no unified definition of Innovation, it is widely acknowledged that innovation is critical factors for the company's economic metrics. Innovative companies are growing faster and earning higher profits (many studies are supporting this argument (Brouwer & Kleinknecht, 1997; Geroski, Machin, & Van Reenen, 1993)). If nevertheless, many companies do not innovate, it is threatened with several types of risks and

uncertainties that cause high levels. For example, only one out of five projects that ever initiated is viable. (Asplund & Sandin, 1999; Cozijnsen, Vrakking, & van IJzerloo, 2000)

Differentiation of Optimization and (disruptive-) Innovation

This thesis focuses on the effects of Big Data analysis of manufacturing machine data on innovation. Therefore, it is somewhat critical to understand the differences in optimization and innovation.

For company representatives filling out the survey understanding the small but significant differences easily, optimization is in this study defined as: *"Introducing incremental improvements to sustain an existing product or service."* Innovation, on the other hand, is defined as: *"Introducing emerging ideas and technologies into an existing category."* Moreover, even more, radical change is disruption: *"Introducing new business models to reach new customer segments"* ("Optimize, Innovate, Disrupt: How They Differ, and Why It Matters," 2016)

Using data, especially machine data for optimization, e.g., process or product optimization is already established and even though still new, quite widely used. One example is McLaren Motors, which use real-time data monitoring of their cars, their competitor's cars and data about the track to optimize their Formula 1 cars. This data is used to adapt and improve the strategy and settings of the vehicle in real time and based on reliable data. (Chandran, 2017)

One of the affordances for this development is the increase in computer power and sophistication of analysis tools. SAP, for example, introduced its S4Hana technology, that uses in-memory calculation, allowing better real-time calculations than traditional database systems.

However, examples like this represent optimization, as an existing product (e.g., the Formula One car) is incrementally improved. Even though the sum of all incremental improvements as well as the application of real-time data analysis itself can be seen as an innovation, there was no source that a newly emerged idea or product was introduced based on the real-time data analysis. As the results are already applied during the races (which is promoted and

therefore documented very well), it seems unlikely that real innovation happens during that short time. Taking into account that the underlying architecture of the car cannot be changed, however, preexisting settings can be changed, and the ideal combination of the conditions and track can be optimized. If the data is also used for innovation and disruption cannot be judged on the currently existing information.

Research Gap

Even though a significant amount of research has been performed already on the individual topics, there is hardly any research on the combination of production, Innovation and Big Data. Furthermore, even though in practice there exist some public cases about the use of Big Data analysis for optimization approaches (e.g., McLaren Motors uses the data generated by its Formula 1 cars, in order to improve the vehicle itself - Sometimes even in real time), no case where this kind of data was used for innovation processes was found.

	Optimization	Innovation
Big Data Analytics	Yes	Yes
Big Data Analytics performed using machine data	Yes	No
Machine Data	Yes	Limited

Figure 1 Comparison of which research has been conducted regarding optimization and innovation

This thesis aims to evaluate the possibilities and already existing usage (if any) of Big Data analysis of production data to trigger innovation. To do that, the thesis tries first to evaluate the preconditions for the further research:

R1: To what extent are manufacturing companies (in Austria, Germany, and Russia) using machine data to perform innovation?

For the analysis the first research question is split into four sub-questions:

1. Are manufacturing companies (in Austria, Germany, and Russia) using the data of the production machines?

2. Are these manufacturing companies using additional sensors to enlarge the data gathered?
3. Is the additional data directed to optimize production only?
4. Do manufacturing companies use the data for innovation processes?

These four questions set the baseline for further research, as only when they are answered at least partially positive, further analysis can be performed. The first part of this thesis is therefore dedicated to finding examples of companies that use machine data for innovation.

As a result of this, the first question is aimed at evaluating the companies' readiness of using production machine data in general. If the data is not used at all yet, there might be problems to integrate the production machine data into the companies' IT system.

The second question is aimed at evaluating the production machines readiness. As the initially built-in sensors of production machines tend to be somewhat production specific, the data generated might be limited.

The third question is aimed at evaluating the additional sensors. In this regard, the new sensors might be production specific as well.

The fourth one is used to gather insights of the production machine data importance in the innovation process.

After that, in the second part of the research the following question is explored:

R2: How can innovation be triggered by Big Data analysis of manufacturing machine data?

For the analysis the research question is again split into three sub-questions:

5. How are these companies linking production machine data and innovation?
6. What were the challenges that companies that implemented Big Data analysis of manufacturing machine data into their innovation process faced?
7. Do these companies have successful innovation based on Big Data analysis of manufacturing machine data yet? If so, in what areas?

With this, the fifth question explores the steps those companies took to link the gathered data and their innovation processes.

The sixth question is aimed to understand the problems that might arise when a company tries to implement Big Data analysis of manufacturing machine data for innovation processes.

The seventh question's target is existing success stories using Big Data analysis of manufacturing machine data.

To answer the research questions, two hypotheses are derived from them and tested:

H₁: Manufacturing machine data is used by companies in Austria, Germany, and Russia.

H₂: Manufacturing companies in Austria, Germany, and Russia are using machine data for innovation as well.

Structure

In the first part of this thesis, the methodology and the reason behind will be explained. In the second part, the survey results will be shown and discussed in detail. Moreover, the statistically significant results will be analyzed. In the third part, the conclusions derived from the results will be displayed and discussed further. The fourth section describes the

limitations regarding the survey and the implications derived from it. Taking all of that into account, the fifth part is of the recommendations and practical suggestions.

Chapter 2: Methodology

Type of research

As there is hardly any academic research existing on the combination of Big Data, innovation and manufacturing companies, there is only the hypothesis, models, and theories of the individual fields available to build upon, but no consolidated basis. This thesis tries to generate a model and assumptions for the combination of these areas. To do that exploratory research is used.

Research Strategy

In the first phase, a survey among manufacturing companies in Austria, Germany, and Russia is performed. This builds the basic understanding if production machine data is used for innovation at all and if so in which form.

In the second phase an exploratory case study with one (or more) of the companies that are using their production machine data for innovation, is performed. The results are used to get a better understanding of the dynamics. This part is the primary source for the answer to the research questions 5 to 7.

The contacts of the manufacturing companies are gathered from the Austrian Industrial Association, the Austrian chamber of commerce, the chamber of commerce in Germany and Russia. However, as data protection laws prohibit most of this associations to forward the survey to their members, additional contact information is gathered through company websites, rankings of the most significant companies of each country, and associations these companies are part of that publish the contact information online.

By this means over 3700 email addresses of companies in Austria, Germany and Russia have been collected and the link to the survey accompanied by an invitation and guidance letter has been sent out to them.

To ensure a high response rate, the survey offered guidance through the questionnaire via phone and provide the results of this paper to the participating companies. This, however, makes anonymization of the individual companies' necessary.

Survey Questions

The questions of the survey are designed to ask about the usage of data in the company in general. This is done for three reasons:

First, to decrease the bias. Innovation and data are two buzzwords. Therefore, the risk exists, that companies might claim to use it when asked for it directly. Thus, the survey is designed to discover, if companies use Big Data analysis of machine data to trigger or execute innovation, indirectly.

Second, this approach allows determining differences in the way different data is used inside the companies.

Third, to motivate companies to participate in the survey, and thereby increase the validity of the data, the results and benchmarks are sent out to the companies that have filled out the survey. Broader data is more valuable to those companies than the very narrow range of machine data used for innovation.

In general, the survey was designed to be short, compact and easy to fill out. The approximate time necessary for answering the 16 Questions plus the part about general information (about the company and the person filling out the survey, in order of being able to perform statistical analysis) was 5-10 minutes. To foster participation, even more, the survey questions, as well as all other information, are available in English, German and Russian. The person filling out the survey can choose the preferred language at the beginning of the study.

Survey

Types of Data being used:

Q1 Which EXTERNAL data is used when developing new products in your company?

(Tick all applicable ones or none if it does not apply)

1. Benchmarks
2. Market research
3. Panel data
4. Other. Please specify:

Q2 Which INTERNAL data is used when developing new products in your company?

(Tick all applicable ones)

1. Financial data
2. Customer data
3. Machine data
4. Quality management data
5. None
6. Other. Please specify:

How is internal data generated?

Q3 Is customization performed to generate data? (Select One)

1. NO, data is only generated if software, machines or processes allow for that without customization.
2. YES, if it is necessary to generate data: software, machines or processes are customized.
3. Other. Please specify:

Q4 What is the goal of data generation? (Tick all applicable ones or none if it does not apply)

1. Data is generated to use it for: Introducing incremental improvements to sustain an existing product or service.

2. Data is generated to use it for: Introducing emerging ideas and technologies into an existing category.
3. Data is generated to use it for: Introducing new business models to reach new customer segments.
4. Other. Please specify:

How is this data stored?

Q5 Ad Hoc Storage: (Tick all applicable ones or none if it does not apply)

1. Data collection starts once the necessity for a specific project is given
2. Data of a specific period in the past is always available
3. Other. Please specify:

Please specify period: ... days

Q6 Regular Storage: (Tick all applicable ones or none if it does not apply)

1. All data is always stored
2. Specific data is always stored
3. Other. Please specify:

Please specify which data is stored:

How is this data examined?

Q7 What are the topics when data analysis is used? (Tick all applicable ones or none if it does not apply)

1. Examination of data is used to test a given hypothesis
2. An analysis is performed based on an open question (e.g., whatever clusters emerge)
3. Other. Please specify:

Q8 Who performs the analysis? (Tick all applicable ones or none if it does not apply)

1. A designated data analytics team performs the analytics.

2. The affected departments perform self-service analytics.
3. Other. Please specify:

Q9 Which techniques are used? (Tick all applicable ones or none if it does not apply)

1. Statistical Analysis
2. Scenarios
3. Clustering
4. Association rule learning
5. Classification Tree Analysis
6. Data fusion and data integration
7. Data Mining
8. Genetic Algorithms
9. Machine Learning
10. Neural Networks
11. Optimization
12. Other. Please specify:

Q10 Who has access? (Tick all applicable ones)

1. Data is accessible for every employee
2. Special data views are accessible to certain departments
3. Only a special data department has access
4. Data is open to customers or suppliers as well
5. Data is open to individual universities
6. Data is open to everyone
7. Other. Please specify:

Please specify the departments with access:

Experiments

Q11 Are experiments performed? (Specific changes in any process, to measure its effects)

1. Yes

2. No

Q12 What kind of experiments? (Tick all applicable ones)

1. Customer-related (E.g., Change of Advertisements or price changes for a part of the customers)
2. Supplier related (E.g., Order amount, Order frequency, etc.)
3. Machine-related (E.g., process cycle duration, process steps, temperature, additives, raw material)
4. Other. Please specify:

Q13 In what environment are experiments performed? (Tick all applicable ones)

1. Regular operations
2. Designated experimental environment
3. Other

Q14 How are experiments triggered? (Tick all applicable ones)

1. Manually
2. Time-based automatized
3. Randomly automatized
4. Other. Please specify:

How does Data help your company?

Q15 How does data help your company concerning PRODUCTS?

	Not at all	Slightly	Somewhat	Very	Extremely	N/A
Make existing products even better	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Create new products	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Reduce costs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q16 How does data help your company concerning CUSTOMERS?

	Not at all	Slightly	Somewhat	Very	Extremely	N/A
Serve existing customers better	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Target new customers	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Survey Design

Question 1 and 2 ask for the type of data being used, to figure out if the company uses machine data at all.

The goal of question 3 is to figure out the relevance of data in the specific company. If the value is high, companies are more willing to undertake the effort of customization, to collect more data.

Question 4 asks for the reason of data collection and thereby gives the first indication if the company uses it for innovation. This question is highly important to distinguish between optimization (which is not the focus of this thesis) and innovation.

Question 5 and 6 have multiple purposes: First the enlarge the data set to provide better insights and benchmarks for the participating companies. Furthermore, the effort of data storage indicates the value of data for the company and helps in evaluating the readiness of the company regarding innovation triggered by data.

Question 7 to 10 elaborate on how the data is used inside the company. The target is to get more indications if machine data is used for innovation and if this is done by using Big Data analytic algorithms.

Question 11 to 14 examine the company's efforts to increase the data and variance of data points by performing experiments. If a company is willing to conduct experiments, this indicates a high readiness for Big Data analytics of machine data. Furthermore, the answers

represent a starting point for the in-depth interviews with promising companies for a case study on how innovation based on Big Data analytics of machine data can be implemented.

Question 15 and 16 test the importance of different types of data to the company. Thereby, the bias is decreased, and the validity of the results on the significance of data for innovation is increased significantly.

All other questions (see appendix) gather general information about the person filling out the survey (e.g., age, gender, position in the company, and so on) and the company (e.g., industry, production company or not, size, and so on).

Interview

After the analysis of the survey results, the companies that are applicable as typical cases to this study are evaluated.

To be applicable, the company must use machine data. Therefore, the Question "Q2 Which INTERNAL data is used when developing new products in your company?" must be answered with machine data (among other possible answers).

Furthermore, the company must perform innovation based on data. Therefore, the Question "Q4 What is the goal of data generation?" must be answered with either "Data is generated to use it for Introducing emerging ideas and technologies into an existing category." or "Data is generated to use it for Introducing new business models to reach new customer segments." or both.

Also, the company (representative) must agree to the interview. A separate question asked for the willingness of the person filling out the survey to take part in a short, in-depth interview. Furthermore, companies that were a fit for the study, but did not indicate the willingness for an interview, were asked again via email and phone.

The interview itself was based on open questions following an interview guideline, that was revised after each interview. Also, clarification questions were asked.

Finale interview guideline

1. Were additional sensors installed in the machines to collect more or different data? If yes, which?
2. Do you connect the machine data with innovation processes, if so how?
Please describe how the process works and which departments are involved? Which data are collected, how are they stored and analyzed or reused?
3. What kind of experiments were performed?
4. Were there any difficulties during the implementation?
5. Was everything done internally, or did you hire external consultancy services?
6. What are the goals/expectations of the company with regards to machine data?
7. Are there already first success stories?

Chapter 3: Survey results

As companies hardly participate in master thesis surveys, a meager response rate was to be expected (especially as information about data is considered sensitive, many companies might deliberately not participate at all). From over 3700 sent out invitations 114 people started the survey, however, only 75 filled out at least one question. In total, the survey was completed by 53 persons. Nevertheless, the survey revealed some interesting details (even though many findings are not significant as the number of responses is too low for any statistical validity)

Findings

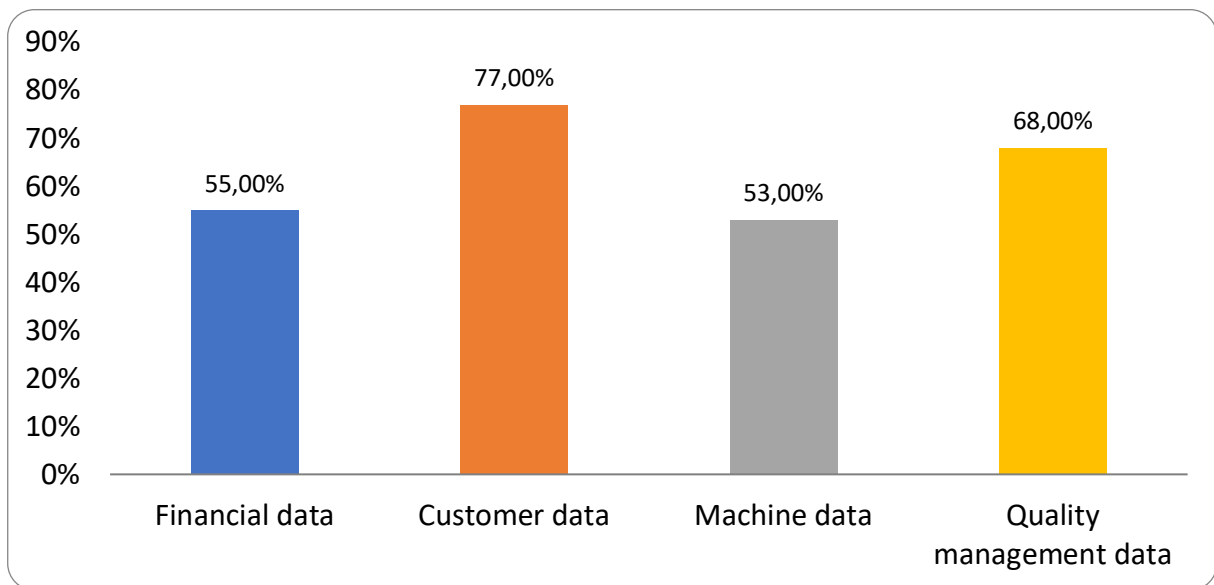


Figure 2 Internal data being used for the creation of new products

When it comes to data being used in the creation process of new products as one would expect customer and quality management data are used most often (77.33% (0.089)¹ Customer data and 68.00% (0.079) quality management data). However, with 53.33% (0.062) machine data is used almost as often as financial data. At least for the surveyed companies that show the increasing value of machine data in product development.

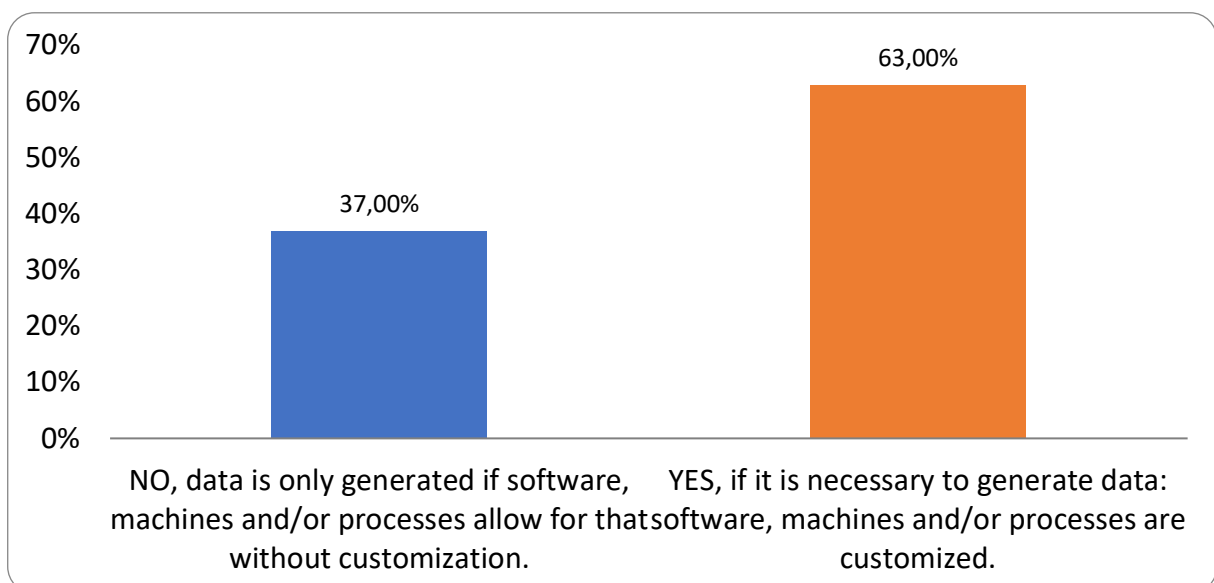


Figure 3 Is customization performed to generate data?

¹ Figures in brackets show the standard error of the mean

Furthermore, the value of data, in general, is seen by the efforts of customization to increase data generation: two-thirds of the surveyed companies are performing customization to generate the needed data. Since not only change costs like additional sensors, reprogramming, and so on are necessary, but also maintenance cost increase (as standard solutions, depending on the amount of customization, might not work), customization can be quite cost intensive. This highly indicates that companies understand the value of data more and more. Also, customization requires specific knowledge, as the company has to understand what additional data is necessary and how to generate that. (see chapter 4 indirect sensors)

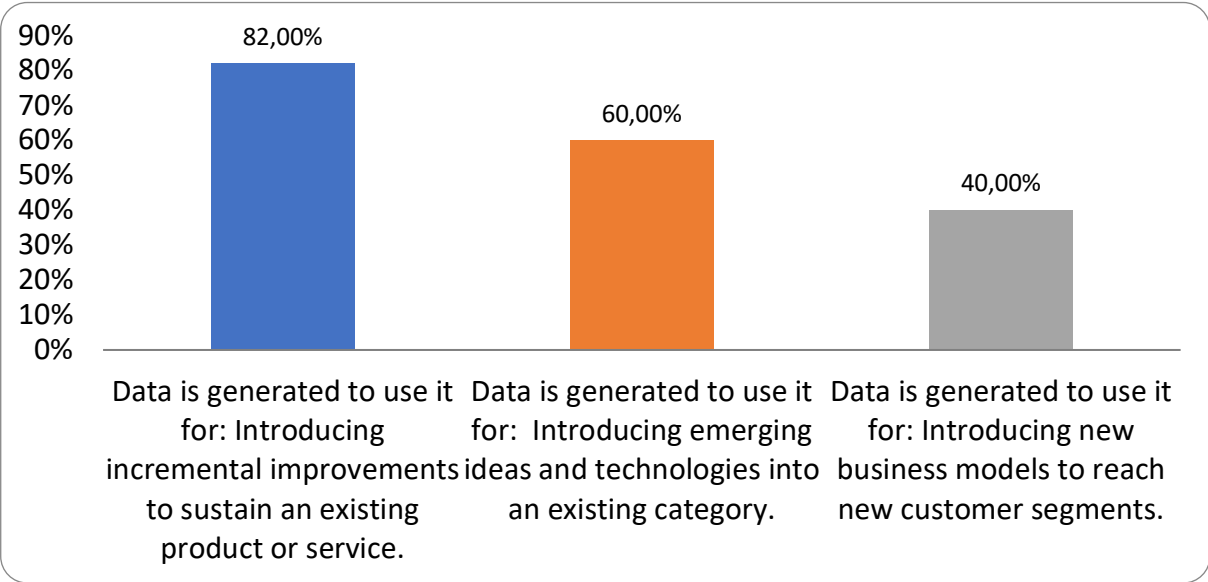


Figure 4 Goal of data generation

According to the survey, 59.68% (0.076) of all companies use data for innovation and 40.32% (0.051) even for disruptive innovation. However, in the interviews, it became clear how hard it was for some companies to differentiate between optimization and innovation, especially when it comes to machine data, only 1 out of the four interviewed companies used this kind of data for their innovation processes. Nevertheless, this figure shows what high expectations companies have regarding data.

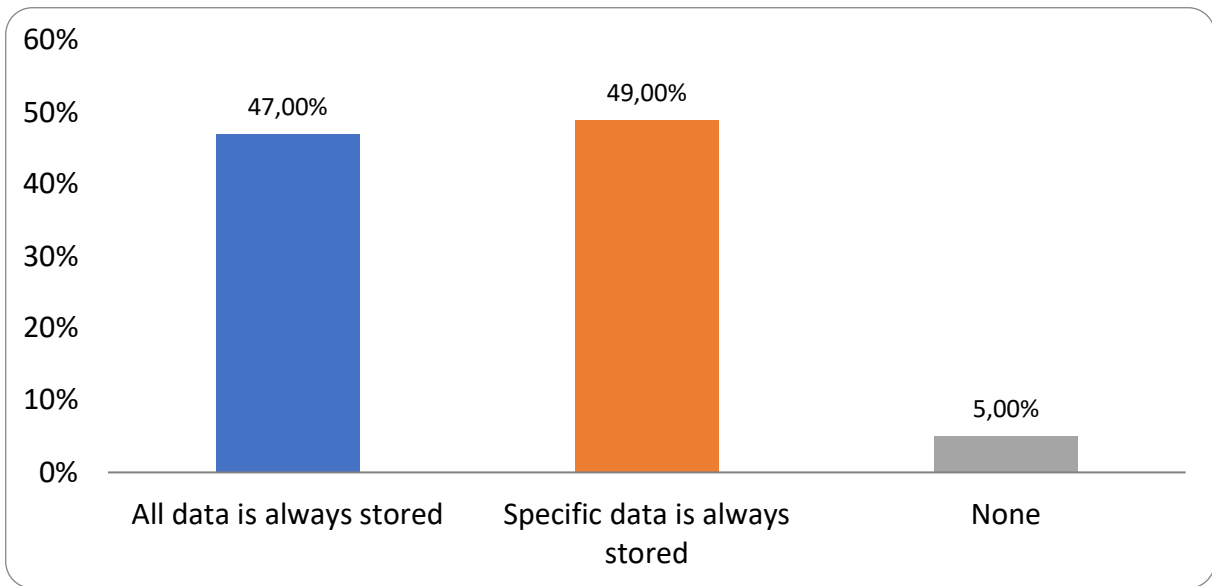


Figure 5 Regular storage

Another indicator that companies highly value data is the regular storage: 47.37% (0.063), almost half of all companies answering that question, store all data permanently, and only 5.26% (0.007) does not store any data at all. At least some specific data, mostly quality management, customer data, financial data, production data or process data, is stored permanently.

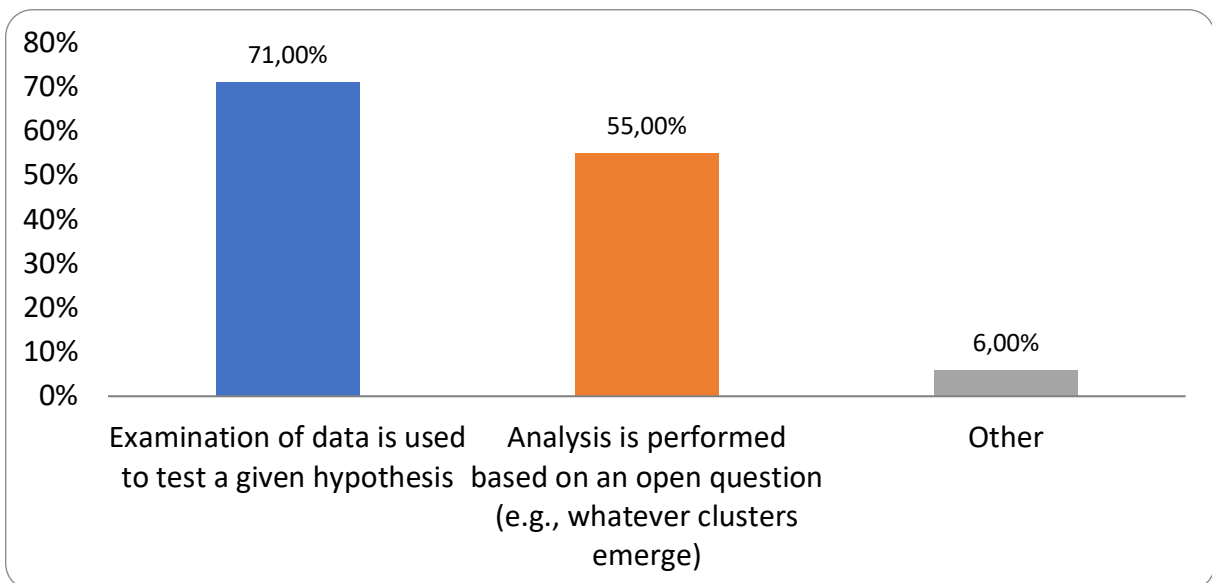


Figure 6 Open or closed analysis of data

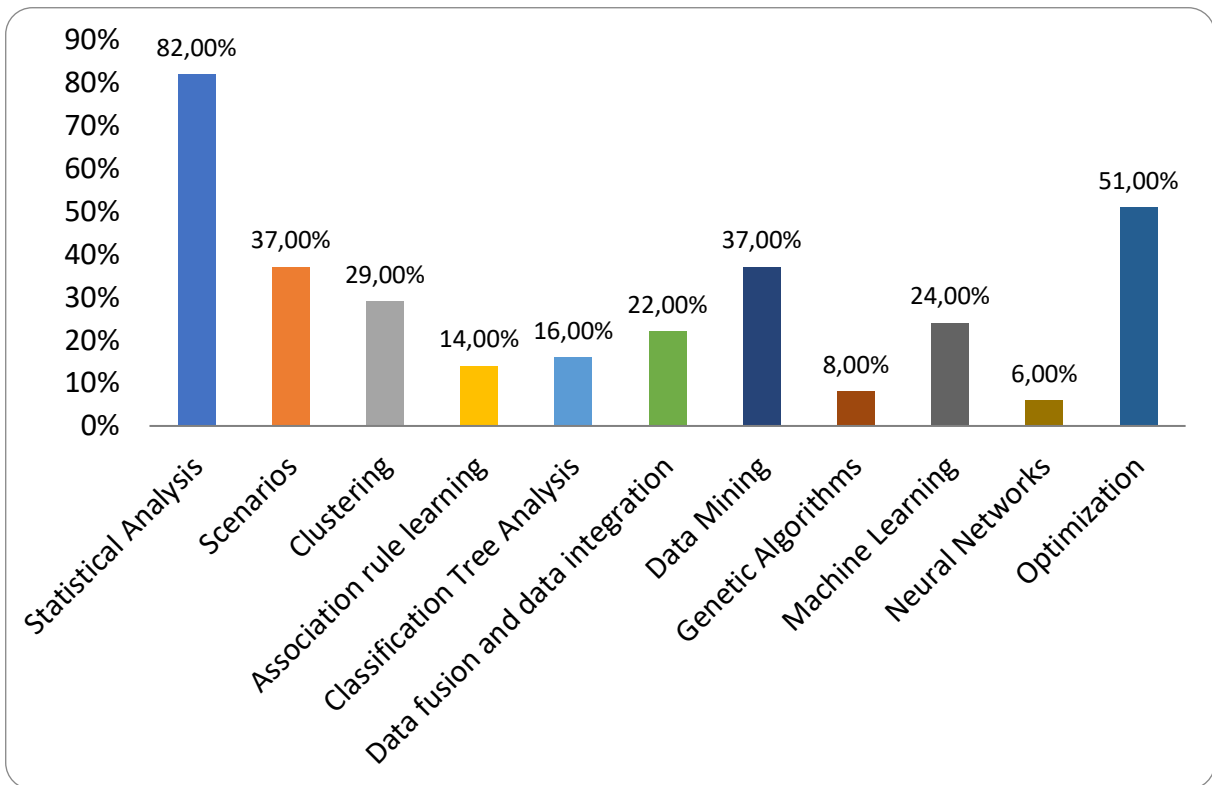


Figure 7 Analysis techniques being used

The approach of data analysis is still mainly about testing a given hypothesis, however Big Data analytic approaches like data mining and clustering are becoming more and more popular among the survey group. Furthermore, even though testing of specific hypothesis is used more often, a broader analysis approach that includes analysis without a given hypothesis, but an open cluster analysis was stated 55.10% (0.079) in the survey.

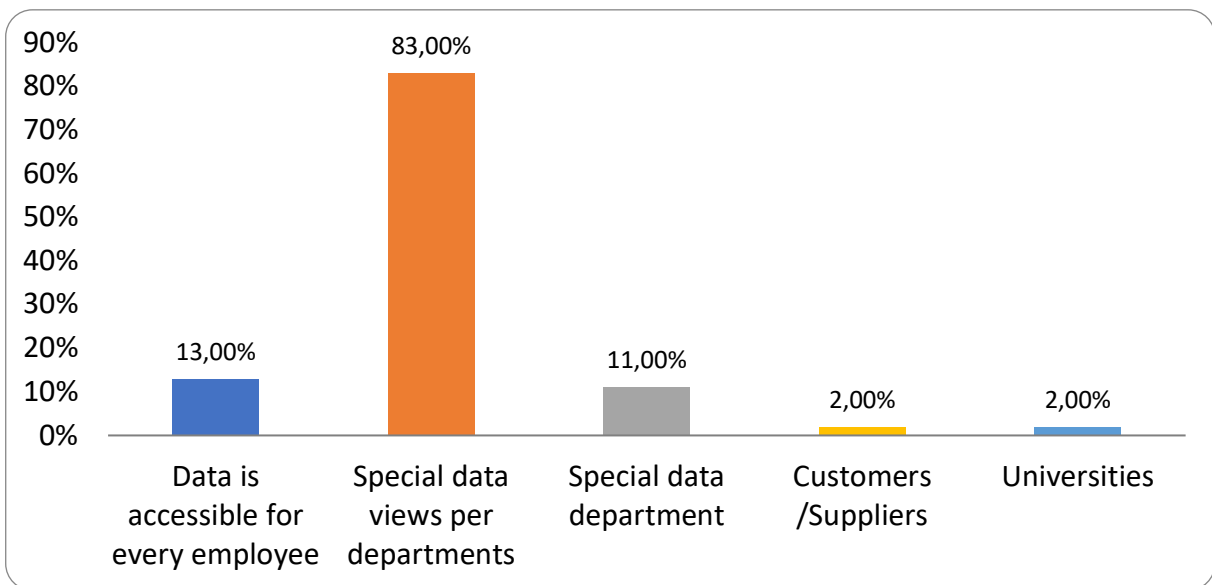


Figure 8 Access to data

In the significant majority of companies, 83.02% (0.114), departments have access to a unique view of the data. This makes sense, as some parts of the data might be sensible information. However, extreme limitations in data views can lead to missed innovation opportunities, as operational, data might help the Research & Development department.

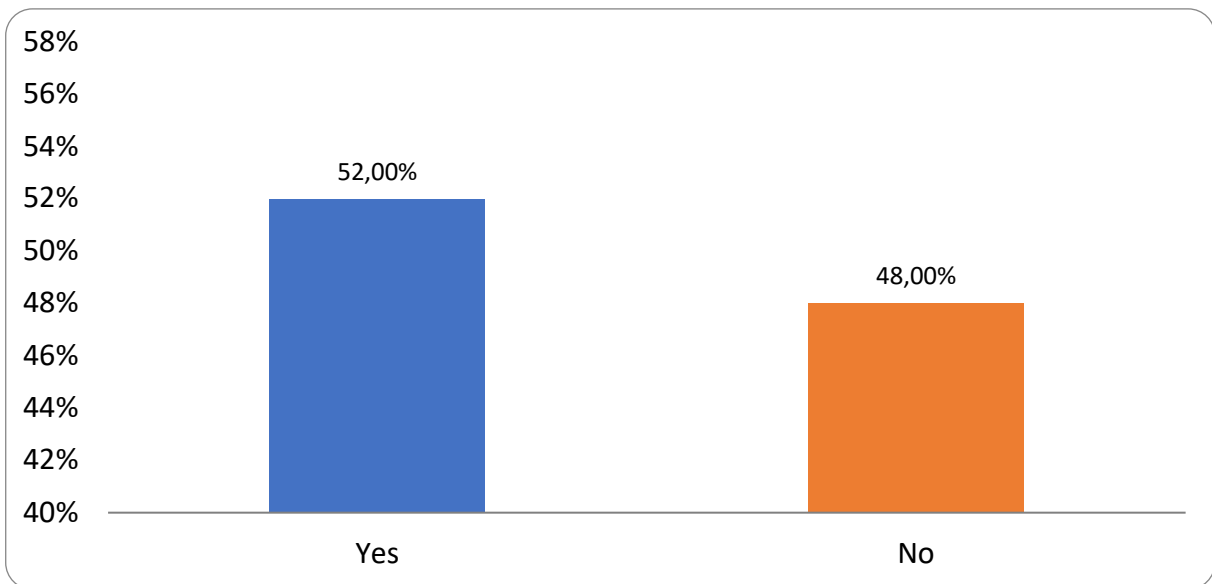


Figure 9 Are experiments performed?

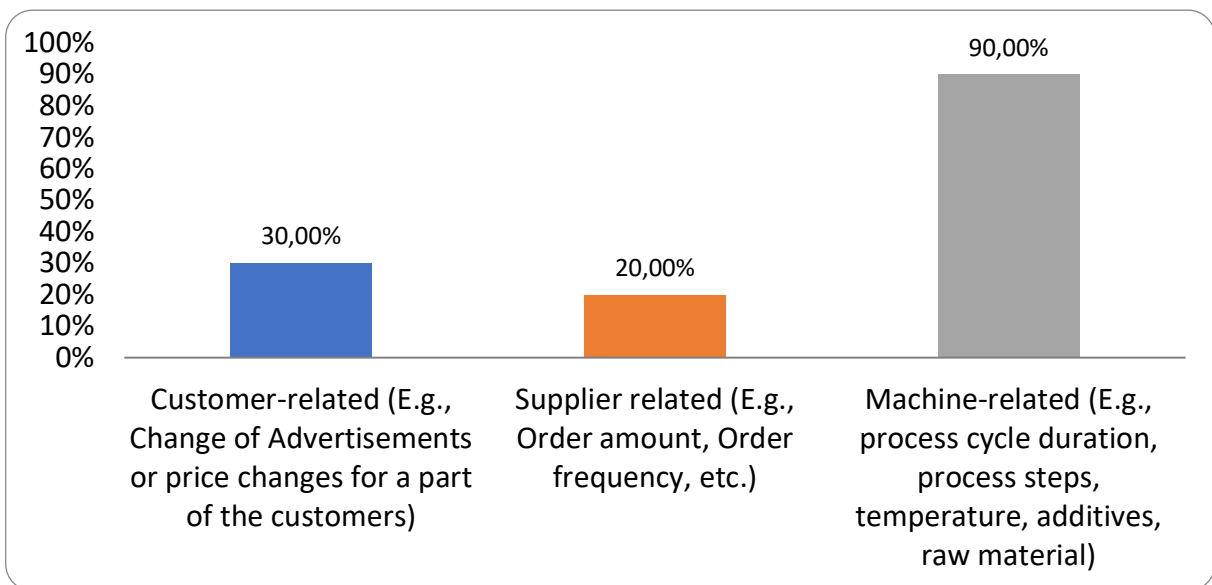


Figure 10 Types of experiments performed.

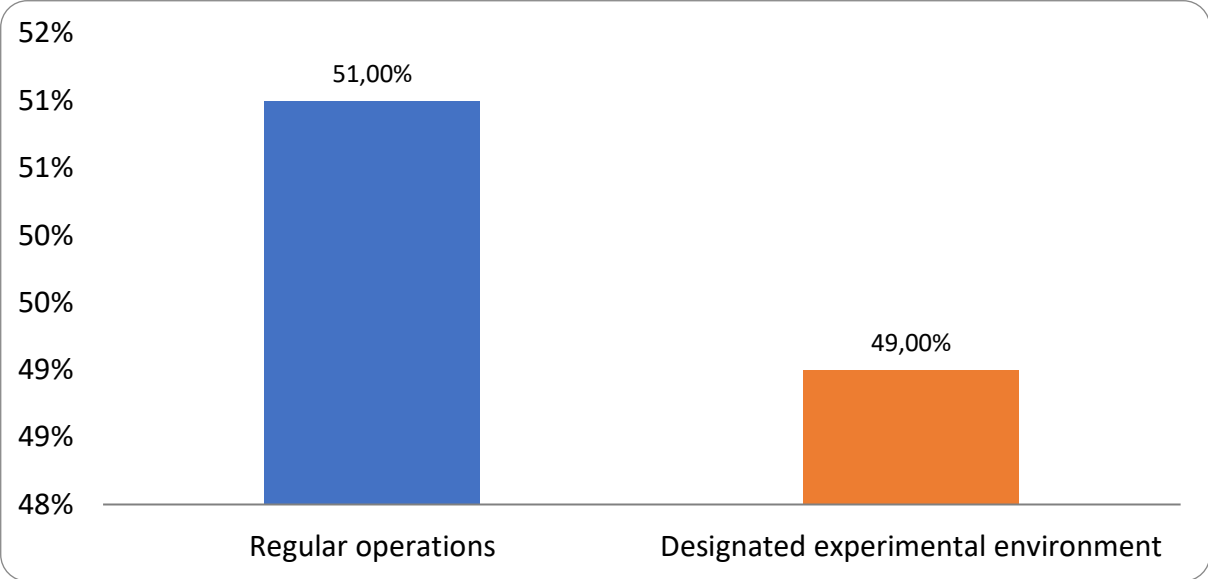


Figure 11 Experiment environment

More than half, 51.72% (SE 0.068), of the companies filling out the survey are using specific changes in processes (experiments) to enlarge the data variety. Almost all of the companies performing experiments are doing this related to machines and machine data (90.00% (SE 0.164)), and even nearly half of the companies are doing it in a specific test environment, even more, 51.16% (SE 0.078), are performing experiments inside the regular operation environment. Doing so enlarges that data variety and creates a much broader data lake that can be used for analysis and in the long run innovation.

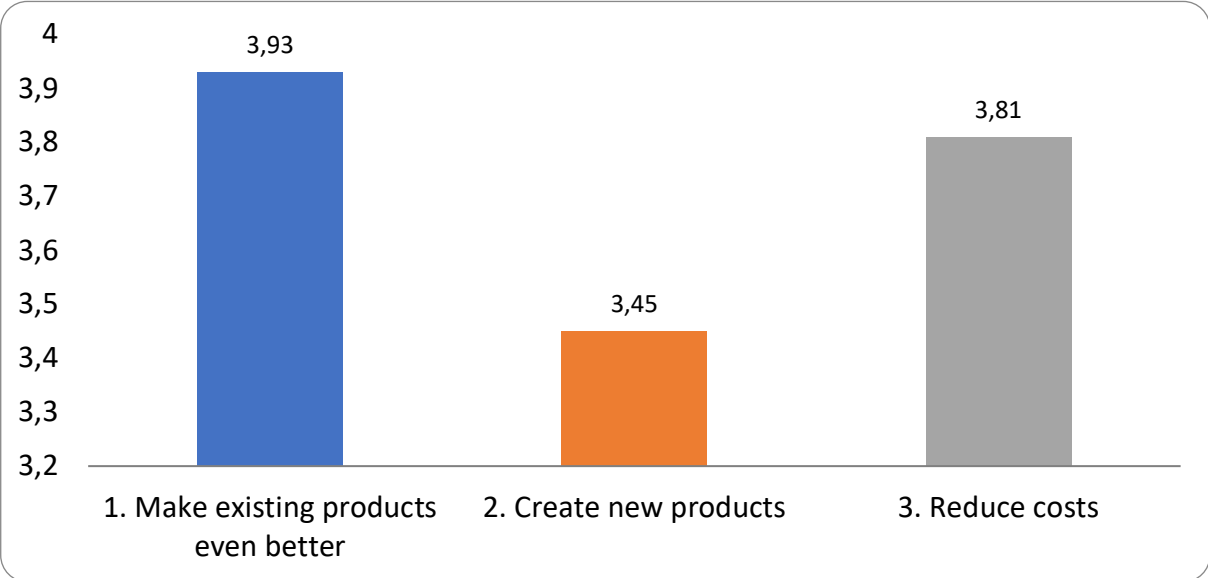


Figure 12 How does data help regarding products?

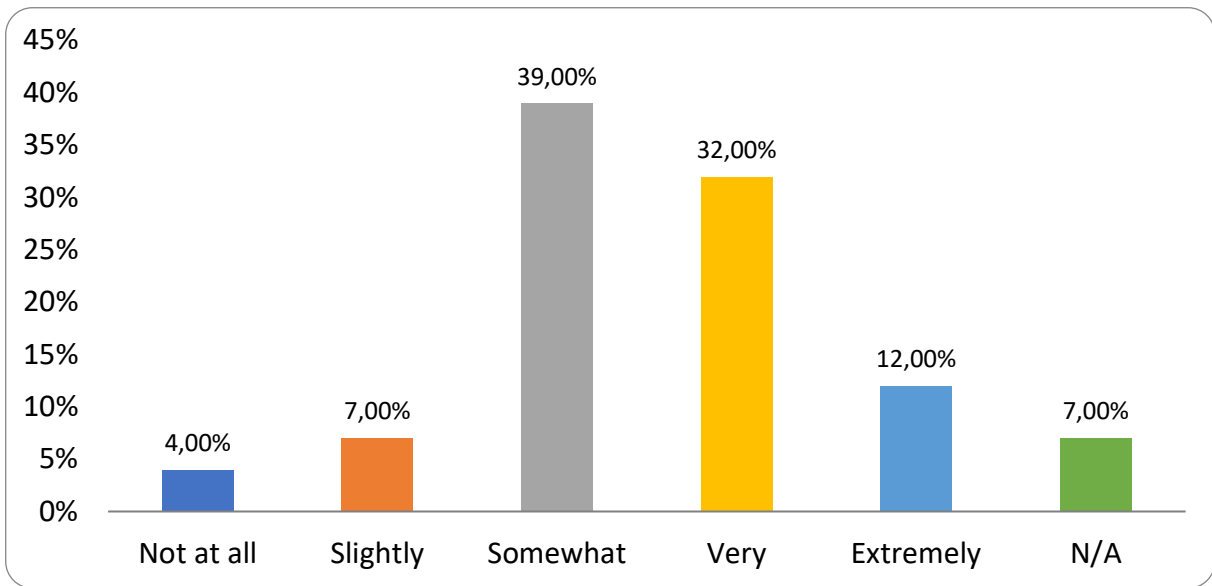


Figure 13 How much does data help regarding creating new products?

The survey question asking how much data helps regarding products shows that the majority assesses the highest benefit from data (and thereby also machine data) is to improve existing products. This is a first indication that most companies focus their data analysis on optimization rather than on innovation. However, due to the limited sample number, the difference is not significant.

Looking into more details how data helps companies creating new products, we can see that 38.60% (SE 0.051) of the companies surveyed estimate the impact of data to be only somewhat influential.

Sample characteristics

The survey was filled out by 75 people and completed by 53. Out of those the majority (47.17% (SE 0.065)) represented companies with 500-10.000 employees. Compared to the number of companies separated by the number of employees in Germany, this constitutes an overrepresentation of such companies, as in reality they are distributed entirely different. (Deutschland, 2018)

< 50 employees	50%
50– 500 employees	46%
> 500 employees	4%

Figure 14 Share of companies by employee size classes

However, as implementing a system that uses machine data for innovation is quite costly, it is by any means improbable to find pioneers in this field among very small companies.

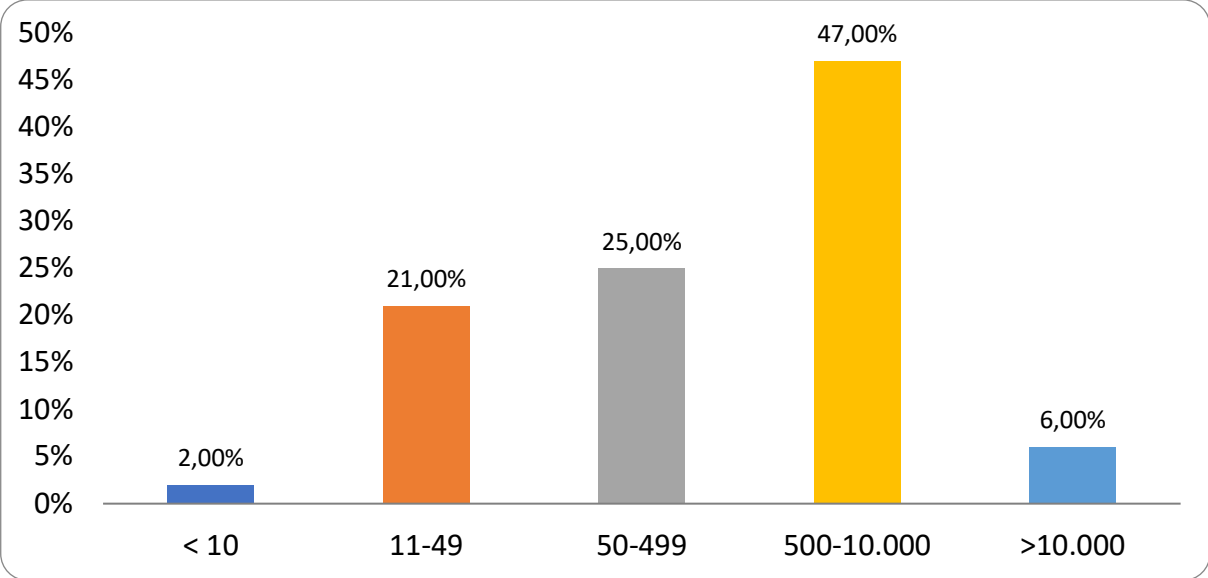


Figure 15 Number of employees

Even though the sample size is rather small, Industries amongst the companies in the survey are vastly spread, which increases the validity of the data. However, regarding geographical location of the headquarters or primary markets there is a massive overrepresentation of Austria (60.78% (SE 0.085)).

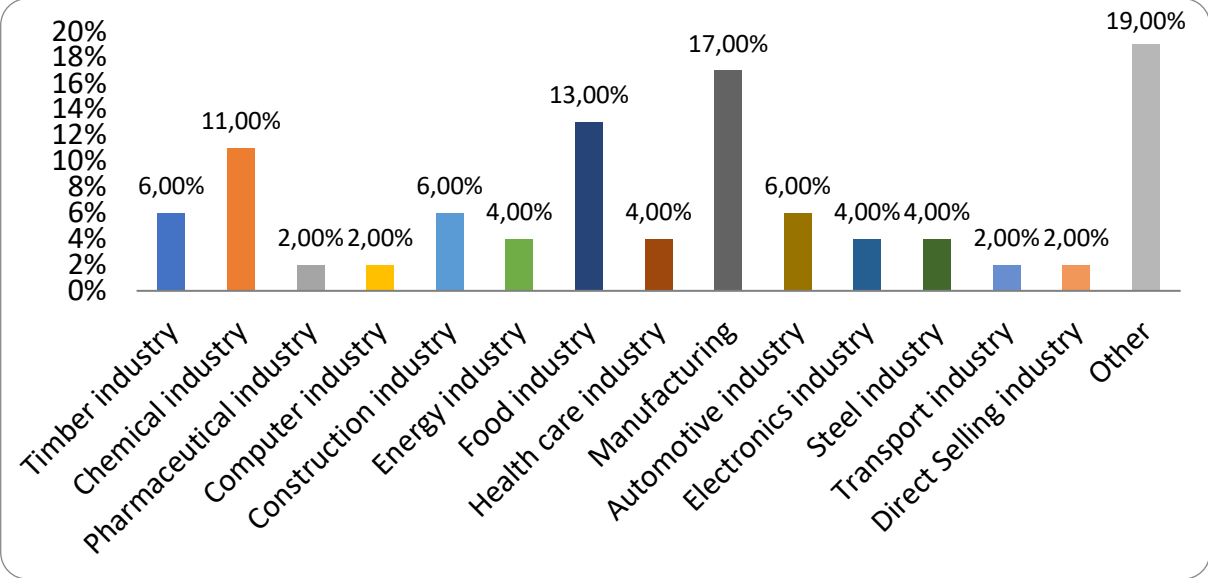


Figure 16 Industry

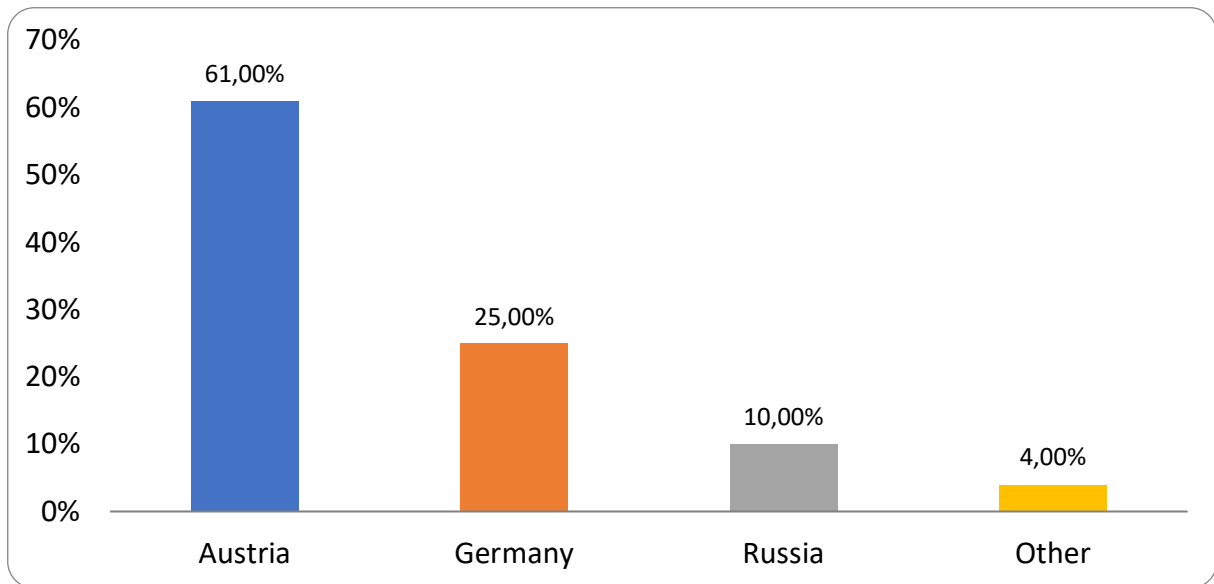


Figure 17 Location of headquarters or major market

Impact of survey results on the research questions

The data of the survey, even though limited in significance due to the small sample size, gives a first indication of the use of machine data for innovation processes in manufacturing companies.

Ranking third, behind quality management and customer data, machine data seems to become increasingly crucial for the development of new products. Especially the willingness to invest in customization and storage portrays how essential machine data and data, in general, are becoming for manufacturing companies.

The eagerness for customization also indicates that companies are willing to enlarge their data gathering by additional sensors. However, this depends on the sense of purpose of this additional data and level of difficulty to generate it (See chapter 4)

Based on the survey itself, many companies seem to use machine data and data in general for innovation processes. However, after the in-depth interviews, it manifested, that for many companies' innovation and optimization are interchangeable synonyms. Therefore, this figure has to be taken with a grain of salt. However, as we will see in chapter 4, some companies are using machine data for innovation processes.

Chapter 4: Multiple Case Study

To understand innovation triggered by machine data analytics better the method of multiple case studies is used.

After analyzing the data, it turned out that five companies are applicable for further interviews. These companies answered to “**Q2** Which INTERNAL data is used when developing new products in your company?” with machine data amongst others. “**Q4** What is the goal of data generation?” was answered with “Data is generated to use it for Introducing emerging ideas and technologies into an existing category.” or “Data is generated to use it for Introducing new business models to reach new customer segments.” or both. Also, these companies agreed to an interview.

The interviews were conducted in total with five people from three different companies. Some of them have been done at the company headquarters itself, the others via a Video call. To create an environment where the company representatives can speak openly about sensible topics (especially how and what data is used, as well as Innovation and R&D) anonymity of the company names as well as the expurgation of highly sensible parts was agreed on. All transcribed interviews can be found in the appendix. As the interviews have been conducted in German, they are also transcribed in German in order not to change the meaning by translation.

Furthermore, the company websites, informational material (internally and externally available ones) and publications were used to enlarge the gathered data and thereby explore the phenomena in greater depth.

Company 1

Processes

Machine data is transferred live from every machine (more than 200 of such globally distributed machines are used in the company) to a particular machine raw data analysis tool. In this tool, the raw data is transformed and loaded into a unique machine database. Furthermore, a part of the data is transferred to the company's ERP system.

For specific problems, e.g., machine malfunction, process errors and product defects, the affected departments analyze the machine data of the affiliated machines. Thereby insights are created by analyzing the data for outliers.

Furthermore, during tool development test machines are created and equipped with additional sensors. The data derived from those test helps to improve the knowledge about the process and is the basis for stabilizing and afterward optimizing the process.

Finally, a designated data insights team uses data analysis tools and algorithms to derive new knowledge from the vast amounts of machine data points.

People

In addition to a particular team that is dedicated to machine data analysis, many other departments, e.g., Research & Development, Tool Construction, Production, and Maintenance are using the data from their specific point of view to improve performance and gather new insights about the processes.

Externals (except suppliers) are only involved as part of a data research cluster in which the company participates.

Issues

The company faced many smaller and more significant issues. Main issues for the company have understood the exact origin of the data points (e.g., the way of measuring, measuring process, calculation method), improving the data analytic skills of the employees and cooperation with machine suppliers.

The first was resolved by internal tests as well as close cooperation with the machine and machine control suppliers. However, even these suppliers were struggling, in the beginning, to answer the precise question regarding sensors and data, as they have not been confronted with such questions before. Therefore, also on the supplier side, the process of data collection in their machine had to be investigated.

The second problem was addressed by both hiring new employees with knowledge in the field of data analysis as well as training existing employees. However, the most significant learnings for both groups of employees happened during the machine data analytic project, as it became clear what questions are essential and what are the most useful tools to answer them.

The third problem was resolved by increasing the level of trust, but also by deluding the questions directed at the suppliers. Moreover, the machine supplier's readiness to answer questions regarding data and measurement grew significantly over time.

Company 2

Processes

Company 2 transfers the data of its machines and by using a smart-box the data of the machines of its customers back to the company. A specially developed software translates the machine data and prepares it for analysis. Before doing that, an engineer had to access the machines directly and copy the data.

The company uses additional sensors to enlarge the data gathered and to receive specific information necessary to understand the processes better. However, this is done only on extraordinary occasions. Typically, the company tries to use the data gathered from sensors and using mathematical models and process knowledge to calculate values that are not measured directly.

In a first step, an experienced engineer has a look at the data. Using his knowledge, he tries to get the first insight. Afterward, a specialized team analyzes the data using analytic data algorithms, partially even Big Data analysis tools. However, this process is in a very early stage and not finalized yet.

People

In addition to the project team, the data and data analytics tools are used by the Research & Development team and the service technicians (e.g., during acceptance of new machines by their customers).

Company 2 includes many different external partners, e.g., for developing and applying unique data analytics algorithms. These external partners are mainly used to leverage their technical knowledge (e.g., universities and software centers). However, the company does not use external partners regarding developing the analysis and innovation process itself. Another essential partner is the company's customers, whom company 2 works very closely with to understand better, how the machines are used at their customer's sites and what issues can arise there.

Issues

One of the significant problems company 2 faced, was to convince its customers to become a partner in this project and share the data of their production machines. Also, to the issue of trust and sharing the benefits equally, also the transmission of the data without directly accessing the customers IT system directly had to be resolved. This was done using a smart box that transmits the data directly from the machine to the company.

Company 3

Processes

In company three all production data is stored in the production control system. Due to the heterogeneity of the IT and even more data environment inside the whole group, the data is then extracted, transformed and loaded into various other data systems. Depending on the given IT landscape available in each business unit the data is transferred from and to a manufacturing execution system and various versions of SAP.

Currently, the focus is on advanced process control, real-time optimization and predictive maintenance. To generate the necessary data, the company uses sensors more extensively, especially for moving and high wear parts. All the sensors used are built by trusted suppliers,

creating specialized sensors for adverse conditions for years. That is necessary as the production conditions the company is operating in are highly regulated concerning safety. However, as the sensors are specially built for these environments, there are no reported issues with those additional sensors either. Moreover, the company focuses on the connectivity between the individual machines of the production, by applying machine learning algorithms the self (real time) optimization is improved vastly.

Right now, only the departments involved in the production have direct access to the data (with some exceptions). To analyze the data and making use of it more efficiently, the company is in currently testing and afterward setting up a platform on which this data can be stored, charted and analyzed by other not directly affected departments as well. The strategy here is, first to have decentralized data analytics in each department with a focus on their internal scope and tasks, in a second step a specialized data task force might be developed. This highly depends on the feasibility and utility as the processes and technologies between the business units vary greatly.

People

As the program is set up in a very decentralized approach, multiple different projects and subprojects are performed to achieve different tasks in the digitalization process. This leads to a high orchestration effort on the one side, however, on the other side, it ensures that the right experts for the different technologies and processes are involved.

To tame the organizational effort, the company included various external partners and consultancies to set up the program right and ensure the use of a best practice approach. This is a novelty for this company, as usually, the company prefers to trust its internal experts rather than hiring vast amounts of external consultants. This shows the complexity of such a program and the expertise necessary to ensure the desired results.

The decentralized approach leads to (sub) projects being performed on a department or business unit level; therefore, the data analysis is being performed by experts on a department level and only as an exception on business unit level let alone on company level. Indeed, this ensures the understanding of the underlying technology; however, synergy

effects from best practices (if not transferred otherwise) and data analytic skills might be lost. Moreover, it bears the risk of further heterogenization due to different approaches and solutions.

Issues

The company faced two main groups of problems. First, it was fundamental to ensure that the necessary experts have sufficient time resources to focus on the digitalization program. Second, the heterogeneity of the data landscape has to be unified, either by aggregation on an additional data layer or by standardization and homogeneity approaches.

Comparison

To understand innovation triggered by machine data analytics better the different examples, which represent best practices, are analyzed comparatively.

In all three cases, the companies choose (at least as part of the full approach) self-service analysis performed by the individual affected departments. As the understanding of the underlying process is imminent for successful data analysis, this approach is relatively reasonable. However, to ensure synergies and knowledge transfer company, one and two introduced a designated data analytics team as well.

Valid for all three companies is the necessity of an adapted data layer to bundle all different data sources and allow for an easy to use analysis and display of the results. However, with increased data heterogeneity and amount of data sources, it becomes more and more complicated to unify the data and set up the right platform. Besides, all three companies faced the challenge how to set up the organizational process to ensure efficient collaboration of various internal and external resources.

Nevertheless, the most significant difference is regarding sensor technology. The challenges to implement additional sensors in the regular operation are highly relying on the production process itself. Therefore, company 1 and two are using other sensor technology

mainly in test equipment, in contrast, company three implements those in the regular production processes on a regular basis.

In the following subchapters the individual phenomena are explored in greater depth:

Additional Sensors

The majority of the interviewed companies do use additional sensors only in exceptional situations. In general, the sensors that are already built in are used. However, the companies try to calculate the missing data, e.g., inside pressure, by using related data, prior knowledge and mathematical models (similar to how the inside pressure of a car tire is calculated by its revolutions per minute and not by an additional sensor inside the car). The data source of this calculations are sensors that are built in by default and are mostly necessary to operate the machine: e.g., pressure and closing force sensors (Interview C2)

However, especially in test environments companies do use additional sensors and measuring devices. As the machines are operating under high temperature and pressure conditions, for company 1, this is the only feasible method to generate additional data and thereby knowledge. Most sensors would not function under these conditions, at least not during regular operations. The company has therefore decided to use the measuring devices in testing tools, which have the same set up as the serial tools, however, are only used, e.g., 100-hour tests. Furthermore, even though some sensor technology can already be used in this harsh environment, as the tools are regularly undergoing maintenance and being assembled and disassembled the initial position of the sensors, and thereby the data validity, will be changed all the time. The primary focus of these tests is to gather additional knowledge about the production process, to understand the influential factors better. However, if necessary to ensure the production quality of a specific part, it would be possible to include some additional sensors and production monitoring systems. (Interview C1a)

To use additional sensors to generate more and more diverse data, first, the durability of the sensors and stability of their set up has to be ensured. Only then can additional sensors be used in regular production operations enabling live data generation and monitoring of other

production process performance indicators. Nowadays, these measuring devices increase the knowledge about the process itself by being used in tests and experiments or at most for a limited time in regular operation machines. However, for some applications in harsh production environments, these affordances were always given. Therefore, the sensors built for such applications are suitable to enlarge the data generated during the regular production. (Interview C3)

Furthermore, the data of other sensors that are available, e.g., building services sensors that measure the production plant temperature and humidity, is merged with the machine data and included in the analysis. (Interview C1b)

Data availability

Internal machines allow for 24/7 online access to the machine data in most of the interviewed companies (Interview C1b & Interview C2). However, an additional challenge is to gather data from the machines being used by the customers. Company 2 that sells injection molding machines managed to convince some of its customers to partner up and use the machine data from regular operations of their customers to enhance the injection molding machines. To gather the data, in contrast to before when a service technician collected it, a so-called smart box transfers the necessary data back to the company. The primary advantage of this set up is that it is not required to access the customer's internal systems and networks, which would be a significant risk and most customers would probably deny the access to the data then. (Interview C2)

Data validity

One of the critical problems in the machine data ventures of the interviewed companies was making sense of the data the machines provided. Primarily when working with different machine models and producers the comparability and reproducibility were hardly ever given. Instead in most cases, it was not apparent how the data was measured and calculated by the machine and machine control. A significant investigation had to be performed to understand the exact source of the data points (Interview C1a & Interview C1b)

Besides, to compare the data of different points in time and even more when comparing different machines, first the processes have to be stabilized to a much greater extent than ever before. Without this stabilization, it is impossible to extract patterns from the data or spot outliers in the dataset. (Interview C1b)

Experiments

All interviewed companies are using the deliberate influence of parameters to test their effects of other parameter and the result in general. It is not only done to enlarge data but also to deliberately evoke known and unknown flaws and defects in the process and the final parts. Understanding what exactly caused a defect, allows the engineers to prevent it in the future by changing, e.g., the tool setup, the construction of the part of the production process. (Interview C1a & Interview C1b & Interview C2)

The majority of these experiments is performed with designated test equipment, however not exclusively. (Interview C2) One of the reasons to opt for a test rather than operation environment is the stability of processes. Especially with highly sensitive production processes a deliberate manipulation of production parameters might be hazardous and pricey. (Interview C3)

One of the main challenges is linking production quality to the exact machine data. The companies that were interviewed have currently only insufficient possibilities in doing so. Quality data is presently mostly available tied to a specific batch, but not more granular than that. To improve that an automatized quality control must be implemented, that generates quality data linked to each part or at least each production cycle. (Interview C1b)

Data analysis

As all the interviewed companies are still in the early stages of using machine data for optimization and innovation, there is currently no standardized approach on how and who analyzes the data. The method all the interviewed companies used was providing access to all departments that might have a use for it (even though sometimes only a limited view of the data is provided). The departments will use the data when faced with challenges and

problems to find solutions based on the provided data. Also, most of the companies have teams that are dedicated to making use of the data and the possibilities that are embedded in the data. (Interview C1b & Interview C2)

The initial analysis itself is usually still heavily depended on the knowledge of application engineers and experts, however especially when multiple factors have to be taken into account, the companies are applying more and more computer-assisted analytic methods and even actual Big Data algorithms. (Interview C1b & Interview C2)

In the analysis, as well as in the setup of structured processes some of the companies are including external partners, especially universities. These external partners, however, usually have a very technical focus and are used to solve technical requirements mainly. In general, all interviewed companies emphasized that the major part of the projects is done internally. Especially, as all the companies assess that the better understanding of their processes based on data analytics will be a sustainable strategic advantage in the future. (Interview C1a & Interview C1b & Interview C2)

The enormous possibilities that machine data provides on the one hand and the costs of running a machine data project, on the other hand, impede the cooperation between companies. Even between customers and suppliers. Customers working on machine data projects in partnership with machine producers fear that all their initial learning, which they were bearing the costs for, might be transferred to other customers of the machine producer. Therefore, it is necessary to create an environment of trust and added value for all stakeholders. (Interview C1b & Interview C2)

Application of machine data

Machine data is used quite diversely in the different companies: Company 1, for example, uses it quite extensively to stabilize and understand their production processes and its powerful levers. Moreover, knowledge about tool performance and construction implications are used to improve the way tools are constructed in the future. In company two machine data is used to predict life duration of the machines or special parts of the machine and generate life duration models. Furthermore, company 2 uses machine data to

understand better how its machines are used at their customer's plants, as similar to cars the results between examination in a laboratory environment might differ vastly from those in the real-world application. (Interview C1a & Interview C1b & Interview C2)

All in all, machine data, at least in the interviewed companies are heavily rooted in research & development optimization and innovation processes. It is clear to the companies that the value of data is tremendous, however, how valuable correctly and what it can be used for precisely is yet to be figured out. One of the challenges in doing so is to convince the rest of the organization, how data can help further improving the company and training them in how to use data analytics for themselves. (Interview C1b)

Machine data triggered innovation.

Even though a high percentage of the survey sample uses its machine data and some also do it under the buzzword "Innovation" the interviews showed that most companies are still only performing optimization with the learnings they gather from analyzing machine data. Several companies see them using machine data at all as a significant innovation, which definitely can be agreed on, as it is a complete game changer for the industry. (Interview C1b & Interview C2)

However, only two examples of innovation triggered by machine data analysis were found among the interviewed companies: Company 2 uses the learning they gather from internal and customer machine data to create life duration models for its machines and critical parts. These models are used to provide their customers with predictive maintenance applications. However, this technology is still very young and is continuously improved and developed further.

The same company also provides intelligent quality measures, that allow for monitoring of the pressure and injection shot weight, which enable holding them perfectly constant over time. With this, the machine data was used to figure out which influence factors have the highest impact on quality.

Chapter 5: Conclusion

The survey, even though having a small sample size, shows some first indications that allow answering most of the baseline research questions, which are aimed at evaluating the use of machine data in general. Also, the interviews shed some more light on the phenomena.

Q1 "Are manufacturing companies (in Austria, Germany, and Russia) using the data of the production machines?". As 53.33% (0.062) of companies surveyed this question is using it, the question can be answered positively. Machine data was the third most mentioned internal source of data for the development of new products.

Q2 "Are these manufacturing companies using additional sensors to enlarge the data gathered?". Based on the survey results on the willingness to customization, 62.69% (0.077), the question could be answered very positively. However, the interviews have shown, that even though additional sensors are used in particular cases, in general, it is much more common to use the sensors that are built in standardized and combine it with mathematical models to calculate missing data. Furthermore, if additional sensors are used, they are more commonly used in test environments and not in the regular production. Mainly as the production environment demands highly durable sensors and the exact positioning of sensors might be affected by maintenance.

Q3 "Is the additional data directed to optimize production only?". Even though the survey shows, that many companies target to use data for innovation (59.68% (0.076)), in the interview only one company could be identified that uses machine data for innovation. This might be partly due to a vague understanding of the differences between innovation and optimization and partly due to the willingness to use it for innovation at some point, but the companies are not there yet. It is still a long way to understand the data and the processes well enough to generate innovation out of machine data.

Q4 "Do manufacturing companies use the data for innovation processes?". Similar to question 3: yes, it is used in innovation processes. However, due to the interviews conducted, most companies are in reality only generating optimization in the end.

Looking at the more in-depth research questions, the interviews conducted with five represents of four companies give us some first insights.

Q5 “How are these companies linking production machine data and innovation?”. Most of the machine data is available live and analyzed by both specialized data teams and departments with specific problems. Even though statistical analysis is still by far the most commonly used analysis method, other methods and even Big Data algorithms are used more and more frequently. The by this means analyzed data is used mainly to understand and improve processes, but in some cases to create new products as well. Furthermore, the interviewed companies used a multi-perspective approach, meaning that not only a dedicated project team analyzed the data, but different departments performed the data analyzes. Each of these departments looks at the data from a unique perspective and to answer different questions. Simultaneously, a dedicated project team is focused on generating value out of data from a more general perspective. This project team works closely with the other departments to align and understand their needs and findings. If this combined effort is orchestrated accordingly, machine data can be successfully used as a trigger for innovation.

Q6 “What were the challenges that companies that implemented Big Data analysis of manufacturing machine data into their innovation process faced?”. The challenges varied from simple understanding how the data is precisely generated, to finding the right people and improving the acceptance of the project inside to organization. However, one of the problems that were mentioned the most was to ensure added value for all participants and environment of trust, when suppliers and customers are working together on a machine data analytics project.

Q7 “Do these companies have successful innovation based on Big Data analysis of manufacturing machine data yet?”. Only one of the interviewed companies had an example of successful organization. This indicates, that even though it seems possible to generate innovation out of machine data, most companies are not equipped with the technology enough yet to do so.

All this information was used to evaluate the two hypotheses and the hypothesis 1, H₁: “Companies use manufacturing machine data.” can be accepted to a full extent.

The second hypothesis “H₂: Manufacturing companies in Austria, Germany, and Russia are using machine data for innovation as well.”, however, is supported only by one example.

Chapter 6: Limitations

As other than this paper, little is yet known about the use of machine data analysis for innovation; this survey is only first exploratory research accompanied by the first set of interviews exploring the phenomena. It gives the first insight into this complex topic, tries to find examples of companies using data for more than just optimization and combines it with real-life examples of organizations trying to make use of machine data. Therefore, it represents a starting point and highlights the motivation, success, and challenges these companies face in using machine data analytics. To explore the issue more deeply, further analysis and interviews have to be conducted.

One of the most significant limitations to the results of this study is the sample size, both in the survey and in the interviews. To increase validity and significance of the data and results, more companies have to be questioned about their use of machine data and innovation processes. However, as it is quite a sensitive subject and companies are hesitant in taking part in such surveys, doing so is not an easy matter.

Another limitation is the geographical focus. To ensure feasibility, the surveyed companies have been limited to Austria, Germany, and Russia, with Austrian companies being highly overrepresented in the survey results (even though the survey has been sent out to many more German companies). Including other parts of Europe and the world in the sample might change the results considerably.

Furthermore, there is a strong bias in the survey that companies that are more advanced in the area of using (machine) data are overrepresented. This is due to two reasons: First, companies that are advanced in this area and the area of digitalization, in general, might be more likely to have a better online presence and might, therefore, be overrepresented in the

mailing lists used to send out the survey invitation to. Second, companies using data more extensively might have a higher incentive to take part in a survey about the use of data. Therefore, a significant portion of companies that are lagging behind in the area of digitalization might not be reached at all or at least not motivated enough to participate in the survey. This would lead to an overly favorable result regarding the use of machine data and innovation.

The differentiation of innovation and optimization is, even though many definitions exist, in the end recurrently a judgment call. Mainly as it highly depends on the framing. Applying a framework that objectively classifies optimization vs. innovation would increase the validity drastically. However, to do so, companies must grant far bigger access to their innovations and optimization projects. Getting this access seems unlikely unless highest anonymity through a large enough sample size can be guaranteed.

Future research should focus on closing these research gaps by increasing the survey and sample size as well as the variety. Also, the progress over time can be monitored. This is especially necessary, as companies are currently just at the beginning of using machine data analytics. Therefore, there will be a steep learning curve. Companies that are not using machine data at all might start soon, companies that only focus on optimization might go one step further and start focusing on innovation and so on.

During the interviews, one topic popped up multiple times: the first mover risk of companies working together with machine producers that might bear the majority of the initial learning costs. Further research could attempt to find and evaluate solutions and mitigations to this problem, e.g., risk sharing contracts (royalties), knowledge transfer systems and joint ventures for machine data analytic projects between machine producers and manufacturing companies. At this moment, the risks and benefits of open innovation could be applied.

Chapter 7: Management implications

Managers should be aware of the value that machine data analytics can bring to a company. (and the survey shows that the importance of the participating companies is already quite high) Every point of data that is produced by production machines increases the value of the

dataset, and the combination and correct alignment (through data analysis techniques) might reveal an entirely new insight.

Already today some companies are using the data produced by their production machines, however, they are not only extracting knowledge from the data itself, but these organizations are increasing their understanding of their production processes by addressing the critical issues, trying to ask the right questions and think about which data points might help them in the future. This knowledge might differentiate successful from unsuccessful companies in the future, as once concrete applications of machine data analytics emerge, it will be those companies that understand data and the processes behind the best, that will be able to implement and make use of those advances first.

A side effect of trying to digitalize the production and analyze production machine data is that to do so; the processes have not only to be understood, but also to be stabilized. As just a stabilized process can reveal its full potential and knowledge when analyzing the data points. Furthermore, a stable process allows for outlier tracking, that can be used both for production control as well as for insight generation of why these outliers appeared. Thereby, an outlier can be the basis of a new optimization or even innovation.

The survey and interviews revealed that even though some companies have already undergone several learning cycles, in general, the industry is just at the beginning and the verge of breakthrough still lies ahead. This gives other companies a chance to catch up, but this does not come without effort. To do so, companies must begin to understand the value of data, how to collect and how to analyze it. Even more important, companies must understand its production processes to an extent where they can transform their knowledge into mathematical data-based models.

Furthermore, the interviews have shown how important it is to convince the whole organization of the value of data and data analytics. If it is only a small research team believing in machine data analytics, the full value of data will never be captured. Data analytics works best when different perspectives are combined both with the data and as well from the analytics side.

Additionally, it is essential to choose the right employees to take part in the machine data analytics project. On the one hand, it is necessary to include experts that know the company and its processes very well, on the other hand as it is an entirely new approach, people with a fresh mind can be helpful to prevent tunnel vision and to be able to see the bigger picture. Moreover, it might be necessary to bring in entirely new knowledge, e.g., data analytics experts, data warehouse experts and (big) data scientists. Nevertheless, this depends highly on the company; some companies might employ employees with these skills already.

A key challenge for managers is to solve the tradeoff of working together with external partners, which sometimes is necessary (e.g., machine producers to understand how data is generated and collected) and sometimes done for the increased efficiency (e.g., research clusters, universities, external consulting). It should not be underestimated how many initial learnings still have to be done and how costly this can be for the organization. (Interview C1b) However, if an organization does not start to gather these learnings right now, it might, of course, be able to copy some from their competitors later, which gives them an advantage as the costs, therefore, are lower. However, there might be much intangible knowledge, especially about processes, that cannot be copied and needs to be built up over time. Lacking this knowledge represents a considerable disadvantage.

Moreover, managers have to evaluate the use of innovation platform. As the link of machine data to the companies' innovation processes involves several internal and external resources, the collaboration has to be structured. Using these collaboration platforms in the right way managers can orchestrate the use of external and internal knowledge to leverage it to understand its processes better and generate innovation. To choose the right platform multiple factors, have to be taken into account. The applicable affordances for a machine data innovation platform must be defined carefully and are highly depending on the company's IT and data landscape as well as the underlying machine data and the company's organizational structure.

In general, even though the sample size is limited the cases should be seen as best practices and give managers the first insight into what is possible, but also what they have to watch out for.

Bibliography

- Albury, D. (2005). Fostering innovation in public services. *Public Money and Management*, 25(1), 51–56.
- Asplund, M., & Sandin, R. (1999). The survival of new products. *Review of Industrial Organization*, 15(3), 219–237.
- Brouwer, E., & Kleinknecht, A. (1997). Measuring the unmeasurable: a country's non-R&D expenditure on product and service innovation. *Research Policy*, 25(8), 1235–1242.
- Chandran, N. H. N. (2017). "The biggest science project on the planet" was on a racetrack on Sunday. CNBC. Retrieved from <https://www.cnbc.com/2017/09/16/big-data-in-formula-one-mclaren-applied-technologies-at-singapore-summit.html>
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From Big Data to big impact. *MIS Quarterly*, 36(4).
- Cozijnsen, A. J., Vrakking, W. J., & van IJzerloo, M. (2000). Success and failure of 50 innovation projects in Dutch companies. *European Journal of Innovation Management*, 3(3), 150–159.
- Cucurull, J., Martí, R., Navarro-Arribas, G., Robles, S., Overeinder, B., & Borrell, J. (2009). Agent mobility architecture based on IEEE-FIPA standards. *Computer Communications*, 32(4), 712–729.
- Cukier, K., & Mayer-Schoenberger, V. (2013). The rise of Big Data: How it's changing the way we think about the world. *Foreign Aff.*, 92, 28.
- Deutschland, S. B. (2018). Beschäftigte und Umsatz der Betriebe im Verarbeitenden Gewerbe: Deutschland, Jahre, Beschäftigtengrößenklassen, Wirtschaftszweige. <https://www-genesis.destatis.de: Destatis. Retrieved from https://www-genesis.destatis.de/genesis/online/logon?sequenz=tabelleErgebnis&selectionname=42271-0005&zeitscheiben=3>
- Dombrowski, U., Riechel, C., & Evers, M. (2014). Industrie 4.0–Die Rolle des Menschen in der vierten industriellen Revolution. *Industrie*, 4, 129–153.
- Drucker, P. (2014). *Innovation and entrepreneurship*. Routledge. Retrieved from <internal-pdf://230.51.72.115/scholar-11.enw>
- Gebler, O. F., Hicks, B., Harrison, A., Barker, M., & Stirling, P. (2016). Towards the implementation of a predictive maintenance strategy: Lessons learned from a case

- study within a waste processing plant. In *European Conference of the Prognostics and Health Management Society 2016* (pp. 1–17).
- Geroski, P., Machin, S., & Van Reenen, J. (1993). The profitability of innovating firms. *The RAND Journal of Economics*, 198–211.
- Hashemian, H. M., & Bean, W. C. (2011). State-of-the-art predictive maintenance techniques. *IEEE Transactions on Instrumentation and Measurement*, 60(10), 3480–3492.
- Jacobs, D., & Snijders, H. (2008). Innovation routine: how managers can support repeated innovation. *Stichting Management Studies*.
- Karre, H., Hammer, M., Kleindienst, M., & Ramsauer, C. (2017). Transition towards an Industry 4.0 state of the LeanLab at Graz University of Technology. *Procedia Manufacturing*, 9, 206–213.
- Matzkovits, J., Saumer, S., Steinbach, F., Zielke, M., & Seitz, J. (2017). Predictive Maintenance–Integration und Kommunikation im Automobilssektor. In *Industrie 4.0* (pp. 83–89). Springer.
- McAfee, A., Brynjolfsson, E., & Davenport, T. H. (2012). Big Data: the management revolution. *Harvard Business Review*, 90(10), 60–68.
- Moore, M., & Hartley, J. (2008). Innovations in governance. *Public Management Review*, 10(1), 3–20.
- Optimize, Innovate, Disrupt: How They Differ, and Why It Matters. (2016). NOBL Collective - The Future of Work. Retrieved from <http://futureofwork.nobl.io/future-of-work/optimize-innovate-disrupt-how-they-differ-and-why-it-matters>
- Payne, R., & Macdonald, B. (2006). Ambient technology—now you see it, now you don't. *Intelligent Spaces*, 199–217.
- Sagiroglu, S., & Sinanc, D. (2013). Big Data: A review. In *Collaboration Technologies and Systems (CTS), 2013 International Conference on* (pp. 42–47). IEEE.
- Skinner, W. (1969). Manufacturing-missing link in corporate strategy.
- Spath, D., Ganschar, O., Gerlach, S., Hämmerle, M., Krause, T., & Schlund, S. (2013). *Produktionsarbeit der Zukunft-Industrie 4.0*. Fraunhofer Verlag Stuttgart.
- Statista.com. (2017). No Title. Retrieved from <https://de.statista.com/statistik/daten/studie/36846/umfrage/anteil-der-wirtschaftsbereiche-am-bruttoinlandsprodukt/>

- Sundmaeker, H., Guillemin, P., Friess, P., & Woelfflé, S. (2010). Vision and challenges for realising the Internet of Things. *Cluster of European Research Projects on the Internet of Things, European Commission*, 3(3), 34–36.
- Swan, M. (2012). Sensor mania! the internet of things, wearable computing, objective metrics, and the quantified self 2.0. *Journal of Sensor and Actuator Networks*, 1(3), 217–253.
- Utterback, J. (1994). Mastering the dynamics of innovation: how companies can seize opportunities in the face of technological change.
- Wang, J., Zhang, L., Duan, L., & Gao, R. X. (2017). A new paradigm of cloud-based predictive maintenance for intelligent manufacturing. *Journal of Intelligent Manufacturing*, 28(5), 1125–1137.
- Wang, Y.-P. E., Lin, X., Adhikary, A., Grovlen, A., Sui, Y., Blankenship, Y., ... Razaghi, H. S. (2017). A primer on 3GPP narrowband Internet of Things. *IEEE Communications Magazine*, 55(3), 117–123.

Appendix

Survey Questions

Types of Data being used.

Which EXTERNAL data is used when developing new products in your company? (Tick all applicable ones or none if it doesn't apply)

1. Benchmarks
2. Market research
3. Panel data
4. None
5. Other. Please specify: _____

Which INTERNAL data is used when developing new products in your company? (Tick all applicable ones)

1. Financial data
2. Customer data
3. Machine data
4. Quality management data
5. None
6. Other. Please specify: _____

How is internal data generated?

Is customization performed in order to generate data? (Select One)

1. NO, data is only generated if software, machines and/or processes allow for that without customization.
2. YES, if it is necessary to generate data: software, machines and/or processes are customized.
3. Other. Please specify: _____

What is the goal of data generation? (Tick all applicable ones or none if it doesn't apply)

1. Data is generated to use it for: Introducing incremental improvements to sustain an existing product or service.
2. Data is generated to use it for: Introducing emerging ideas and technologies into an existing category.
3. Data is generated to use it for: Introducing new business models to reach new customer segments.
4. Other. Please specify: _____

How is this data stored?

Ad Hoc Storage (Event based storage) (Tick all applicable ones or none if it doesn't apply)

1. Data collection starts once necessity for a specific project is given
2. Data of a specific period in the past is always available (e.g. last 30 days of Quality Management data is always stored)

3. None
4. Other. Please specify: _____

Please specify period: ... days

Regular Storage:(Tick all applicable ones or none if it doesn't apply)

1. All data is always stored
2. Specific data is always stored
3. None
4. Other. Please specify: _____

Please specify which data is stored:

How is this data examined?

What are the topics when data analysis is used? (Tick all applicable ones or none if it doesn't apply)

1. Examination of data is used to test a given hypothesis
2. Analysis is performed based on an open question (e.g., whatever clusters emerge)
3. Other. Please specify: _____

Who performs the analysis? (Tick all applicable ones or none if it doesn't apply)

1. A designated data analytics team performs the analytics.
2. Self-service analytics is performed by the affected departments.
3. Other. Please specify: _____

Which techniques are used? (Tick all applicable ones or none if it doesn't apply)

1. Statistical Analysis
2. Scenarios
3. Clustering
4. Association rule learning
5. Classification Tree Analysis
6. Data fusion and data integration
7. Data Mining
8. Genetic Algorithms
9. Machine Learning
10. Neural Networks
11. Optimization

12. Other. Please specify _____

Who has access?(Tick all applicable ones)

1. Data is accessible for every employee
2. Special data views are accessible to certain departments
3. Only a special data department has access
4. Data is open to customers and/or suppliers as well
5. Data is open to individual universities
6. Data is open to everyone
7. Other. Please specify:

Please specify the departments with access:

Experiments

Are experiments performed? (Specific changes in any process, to measure its effects)

1. Yes
2. No

What kind of experiments? (Tick all applicable ones)

1. Customer-related (E.g., Change of Advertisements or price changes for a part of the customers)
2. Supplier related (E.g., Order amount, Order frequency, etc.)
3. Machine-related (E.g., process cycle duration, process steps, temperature, additives, raw material)
4. Other. Please specify:

In what environment are experiments performed? (Tick all applicable ones)

1. Regular operations
2. Designated experimental environment
3. Other

How are experiments triggered? (Tick all applicable ones)

1. Manually
2. Time based automatized
3. Randomly automatized
4. Other. Please specify: _____

How does Data help your company?

How does data help your company concerning PRODUCTS?

	Not at all	Slightly	Somewhat	Very	Extremely	N/A
Make existing products even better	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Create new products	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Reduce costs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

How does data help your company concerning CUSTOMERS?

	Not at all	Slightly	Somewhat	Very	Extremely	N/A
Serve existing customers better	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Target new customers	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

General Information

If you choose not to reveal this information, please leave the question blank

Company:

Number of Employees

1. < 10
2. 11-49
3. 50-499
4. 500-10.000
5. >10.000

Industry

1. Aerospace industry
2. Agriculture
3. Fishing industry
4. Timber industry
5. Tobacco industry
6. Chemical industry
7. Pharmaceutical industry
8. Computer industry
9. Construction industry
10. Energy industry
11. Food industry
12. Health care industry
13. Information industry
14. Manufacturing
15. Automotive industry
16. Electronics industry
17. Pulp and paper industry
18. Steel industry
19. Shipbuilding industry
20. Mass media

21. Mining
22. Telecommunications industry
23. Transport industry
24. Water industry
25. Direct Selling industry
26. Other. Please specify: _____

Location of Headquarters or major market

1. Austria
2. Germany
3. Russia
4. Other. Please specify: _____

Name of the company

May the company information given be stated in the thesis?

1. 1. Yes
2. 2. No
3. 3. Only if the company name is anonymized
4. 4. Other. Please specify: _____

Is the company at least partially a PRODUCTION COMPANY?

1. 1. Yes
2. 2. No

The person filling out this survey

Gender

1. 1. Female
2. 2. Male
3. 3. I prefer not to identify
4. 4. Other _____

Age in years

Position in the company

Would you be available for a SHORT INTERVIEW about the use of data in your company?

1. 1. Yes
2. 2. No

First Name

Last Name

Phone

Email Address

Are you interested in the results?

1. 1. Yes
2. 2. No

Email Address for the results



Interview with interview partner A from Company 1 (Interview C1a)

Einleitung: Begrüßung und nochmaliges Spezifikation der Unterschiede zwischen Optimierung und Innovation

S: Wofür werden Maschinen Daten im Unternehmen eingesetzt?

I: Daten werden einerseits für die Planung benötigt, dass wir Daten abrufen können von einem Leitstand um zu sehen wo die Maschine aktuell steht, in welchem Schritt des Prozesses. Es ist eine Kontrolle zwischen der Planung und der Fertigung selbst. Hier sind wir eher weniger in dem Bereich neue Produkte zu generieren. Ganz falsch ist der Ansatz aber trotzdem nicht, denn wir versuchen über Berechnungen und Simulationen, Rückschlüsse zu ziehen und mit der Praxis zu verbinden. Zurzeit werden die Probleme meistens im Nachhinein festgestellt und dann versucht die Herkunft dieser zu simulieren. Und aus diesen Erkenntnissen dann Ableitungen findet und z.b. an der Geometrie, Wärmehaushalt etc. etwas verändert. Uns geht es hierbei auch gar nicht zu sehr auf das Produkt selbst sondern mehr um die Fertigung des Werkzeugs und natürlich auch die Prozesse zu überwachen um zu sehen welche Fehler auftreten können. Ziel ist es mit den Erkenntnissen diese in Zukunft zu vermeiden und wenn diese dennoch vorkommen, noch besser darauf zu reagieren.

S: In zukünftigen Werkzeugen wird dieses Wissen dann verwendet und vielleicht sogar anders konstruiert um diese Probleme zu vermeiden, ist das richtig?

I: Ja genau. Das Thema ist generell Industrie 4.0, wir machen das nicht nur weil es modern ist, sondern es ist einfach vieles an Daten vorhanden und wir versuchen das ganze transparenter aufzubereiten. Wir versuchen, dass wir wegkommen von Eingaben die durch Menschen getätigt werden, hin zu einer automatischen Eingabe dieser Daten durch die Maschine selbst. Z.b. Prozessdauer, Zeitpunkt wann der Prozess beendet wurde, Zeitpunkt wann dieser wieder gestartet wurde etc. Das eingeben dieser Daten wäre für einen

Menschen zu mühsam und bringt keinen Mehrwert für eine Maschine ist es in Sekunden erledigt. Zusätzlich birgt die Eingabe durch Menschen eine gewisse Fehleranfälligkeit. Bei einer menschlichen Eingabe muss auch beachtet werden das die Art und Weise wie die Daten erfasst werden variieren könnten, bei einer Maschine können wir das standardisieren und genauso erfassen, aufbereiten und weiterverarbeiten wie wir als Unternehmen uns das vorstellen. Neben Daten zu Zeitpunkten und Dauer messen wir auch Verschleiß an den Maschinen, um darüber Kenntnis zu erhalten wann wir z.b. eine Schwester Maschine einsetzen müssen. Bei diesen Messungen lernen wir immer wieder Neues und die Daten bilden die Basis für Entscheidungskriterien bei neuen Herstellern z.b. oder neuen Prozessen. Die Vergleichbarkeit wird dadurch entschieden erhöht. In der Wiederbeschaffung, Einkauf, nutzen wir das als Entscheidungsgrundlage. Zurzeit sind wir mitten im Projekt, dass wir versuchen alle diese Daten absaugen zu können.

S: Das heißt es gibt noch keinen standardisierten Prozess wie mit Daten und deren Analyse umgegangen wird? Das heißt bei einem Problemfall wird in die Daten geschaut und dann analysiert was das Problem sein könnte?

I: Genau, gewisse Daten werden auch manuell dazu erfasst in dem sie auf den Auftrag gebucht werden. Wir wollen uns aber auch gewisse Alarme z.b. wie oft steht die Maschine aus den Daten herausholen. Damit man das dann filtern kann und auch irgendwann im Gespräch mit den Maschinenherstellern darauf hinweisen kann, dass diese Probleme regelmäßig auftreten und gemeinsam nach einer Lösung sucht diese zu beheben oder gar zu vermeiden. Das wird natürlich auch jetzt schon gemacht allerdings fehlen oft die Daten wie oft und wann genau die Probleme auftreten. Hier kann man die Daten direkt heranziehen und sagen wo es großes Verbesserungspotential für die Zukunft gibt. Damit soll sichergestellt werden, dass die nächste Maschinengeneration dieses Problem nicht mehr aufweist oder der Hersteller eine gewisse Unterstützung in diesem Bereich anbieten kann. Aber das aktuelle Hauptthema ist wie gesagt die Planung, damit wir hier die Maschinen möglichst zu 100% ausnutzen können. Oder wir zumindest den 100% immer näherkommen. Der Planungshorizont wird dadurch deutlich erhöht und wir können weit vorher schon sagen wie weit wir zu einem bestimmten Zeitpunkt voraussichtlich ausgelastet sind. Und dadurch haben wir auch immer einen Abgleich mit dem wie es wirklich aussieht da z.b. Stillstands

Zeiten Unterbrechungen oder andere Arbeiten vorkommen können. Das ist für uns das Hauptthema mit dem wir vom Werkzeugbau an das Thema Maschinendatenerfassung herangehen.

S: Sind extra Sensoren die in der Maschine und Werkzeug verbaut werden geplant oder bereits im Einsatz?

I: In der Maschine weniger, aber in Werkzeugen haben wir bereits Test gemacht Um mehr wissen daraus zu generieren. Z.b. welche Drücke, welche Temperaturen haben wir in gewissen Bereichen. Und was passiert, wenn wir mit diesen Werten herumspielen.

S: Sind diese Versuche in Testumgebungen oder wirklich im operativen Betrieb durchgeführt worden?

I: Die Werkzeuge sind Testwerkzeuge, also nicht wirklich Produktionswerkzeuge. Allerdings sind diese genauso aufgebaut ansonsten könnte man die Erkenntnisse ja nicht für die Produktion verwenden. Da eine Vergleichbarkeit nicht gewährleistet wäre. Wäre der Aufbau unterschiedlich, wären die Daten die wir aus zusätzlichen Sensoren generieren möglicherweise verfälscht.

S: Wäre es auch möglich in Serien Werkzeugen z.b. Einflussgeschwindigkeiten oder gewisse Drücke zu messen?

I: Grundsätzlich wäre es möglich, zurzeit wird es aber nicht eingesetzt. Im Bereich Sensorik gibt es hier die Herausforderung, dass unsere Werkzeuge sehr heiß sind. Und alles mit hoher Temperatur z.b. über 200 Grad ist im Bereich der Sensorik natürlich etwas schwieriger. Hohe Drücke führen zu ähnlichen Herausforderungen, zumindest im Dauerbetrieb, im Vergleich zu einem Test der deutlich kürzer andauert. Z.b. 100 Stunden. Wir haben hier Testserien durchgeführt, allerdings nicht mit dem Fokus die Werkzeuge im operativen Betrieb zu überwachen, sondern um herauszufinden was spielt sich im Werkzeug selbst genau ab. Es ist allerdings nicht ausgeschlossen, dass wenn es z.b. das Produkt fordert wir in der Zukunft

auch eine derartige Sensorik und Überwachung einbauen werden. Allerdings ist hier noch nichts fix geplant.

Interview with interview partner B from Company 1 (Interview C1b)

Einleitung: Begrüßung und nochmaliges Spezifikation der Unterschiede zwischen Optimierung und Innovation

S: Wie werden die Daten von der Maschine abgeholt bzw. wie läuft Zugriff ab. Werden die Daten zum Beispiel einmal am Tag abgeholt oder gibt es eine Live Verbindung und zusätzlich wäre interessant ob es zusätzliche verbaute Sensoren in der Maschine gibt?

I: Maschinen sind ständig online und es werden auch alle bzw. fast alle Daten die die Maschine aufzeichnet verwendet. Allerdings denken wir, dass die Maschinen je nach Maschinensteuerung Daten mit unterschiedlich hohen Informationsgehalt zur Verfügung stellen. Für uns ist es das ein sehr wichtiger Punkt, da man als Endnutzer nicht weiß wie die Daten im Detail generiert wurden. Um allerdings die notwendige Qualität bei der Auswertung der Daten zu erreichen, muss eruiert werden wie die Daten zustande gekommen sind. Ansonsten kann nicht ermittelt werden welche Aussagegrad die Daten bieten.

S: Ist der Lösungsansatz des Unternehmens hierbei mit eigenen Sensoren und eigener Datenerfassung vorzugehen?

I: Der Ansatz des Unternehmens ist mit den Maschinenherstellern zusammenzuarbeiten um gemeinsam herauszufinden wie die Daten im Detail generiert wurden. Dieser Ansatz wird vor allem deshalb gewählt da es bei dem bereits existierend Maschinenpark keinen Sinn machen würde in jeder Maschine neue zusätzliche Sensoren zu verbauen um Daten zu generieren die ohnehin bereits vorhanden sind. Wir wollen vor allem die Daten nutzen die bereits vorhanden sind. Der Grenzgang hier ist allerdings, dass wir um zu verstehen wie die Daten generiert zu werden die Maschinenhersteller Fragen stellen müssen. Diese Fragestellung gibt natürlich auch dem Maschinenhersteller selbst ein gewisses wissen, das auch mit anderen Kunden teilen könnte.

S: Das heißt es handelt sich hier auch um ein strategisches Problem, wie man bestmöglich mit dem Maschinenhersteller zusammenarbeitet, ohne für eventuelle Konkurrenten eine kostenlose Vorarbeit zu leisten?

I: Ganz genau. Wir haben zum Beispiel in einem Vorprojekt, die gesamte Hardware einer Maschine ausgewertet und damit den gesamten Prozess vermessen. Parallel zu dem was in der Maschine schon vorhanden ist. Ziel hierbei war es alles so modular aufzubauen, dass wir mit den Erkenntnissen zu jeder beliebigen Maschine gehen können und dort alle messtechnischen Fragestellungen beantworten können. Da natürlich manche Maschinen die notwendigen Sensoren dafür noch nicht eingebaut haben, ist das für uns ein Lerneffekt um zu eruieren was in einer neuen Maschinengeneration vorhanden sein sollte. In manchen Fällen werden zusätzliche ältere Maschinen Generationen von uns mit den notwendigen Sensoren ausgerüstet.

S: Wird in den Teilen der Maschine die vom Unternehmen selbst hergestellt werden zusätzliche Sensorik verbaut? In einem vorangegangenen Interview wurde mir dazu erklärt das besonders die hohen Drücke und die hohe Temperatur Probleme bereiten und daher Sensorik an zwar in einem Testbetrieb verbaut werden kann, allerdings in der eigentlichen Produktion dies kaum möglich ist.

I: Im Bereich der firmeneigenen hergestellten Teile der Maschine haben wir eine große Fragestellung und zwar wie gehen wir mit dem Material aus dem die Teile geformt werden um. Die Viskosität des Materials ist im Lagerzustand eine völlig andere als in der Maschine selbst. In der Maschine verhält sich das Material sehr dünnflüssig. Wir brauchen aber im Werkzeug einen Spalt zur Entlüftung. Wenn er zu groß ist, kann Material in den Spalt eindringen und Messwerte verändern. Damit würde sich der Messwert während dem Prozess ständig ändern daher wird Sensorik in diesem Fall vor allem im Laborbereich eingesetzt und nicht in der eigentlichen Produktion. Denn im Laborbereich können diese Veränderungen der Messwerte sehr genau beobachtet werden und auch immer wieder auf Plausibilität geprüft werden. In der Produktion wäre dies nicht möglich. Es gibt hierzu zwar schon einige Fortschritte, vor allem im Bereich der Temperatur Sensorik allerdings ist es in der Größenordnung in der wir produzieren nicht mehr möglich, jede einzelne Maschine mit

einer solchen Messvorrichtung auszustatten. Zusätzlich muss auch beachtet werden, dass die Sensoren sensible Messgeräte sind, die nicht unbedingt für die raue Produktionsumgebung geeignet sind. z.B. werden immer mal wieder Wartungsarbeiten durchgeführt, bei denen Teile der Maschine auseinander danach wieder zusammengebaut werden. Damit wären die Messgeräte ständig Veränderungen ausgesetzt die die Ergebnisse verzerren würden.

S: Werden von der Maschine auch Daten über die Umgebung z.B. Temperatur Luftfeuchtigkeit gemessen?

I: Nein das wird von der Maschine nicht gemacht hier greifen wir auf die Daten von anderen Datenbanken zu. Z.B. auf die Daten der Haustechnik.

S: Grundsätzlich noch einmal zur Auswertung der Daten wie vorhin besprochen sind die Daten einer Maschine ständig live, wie werden diese gespeichert und von wem und in welcher Art und Weise werden diese analysiert?

Es wurde beim Maschinenhersteller eine Software gekauft in diese werden alle Daten gespeist und gewisse relevante Daten werden im ERP-System gespeichert. Zusätzlich werden manche Daten auch in einer extra Datenbank gespeichert.

Die Daten werden je nach Fragestellung ausgewertet. Da wir mit allen Daten immer live sind gibt es auch ganz unterschiedliche Herangehensweisen, Beispielsweise kann sich ein Maschineninstandhalter anschauen wie die Maschine aktuell läuft. Oder auch ein Anwendungstechniker der wissen möchte wie eine Maschine im Vergleich zu einer anderen läuft. Wir haben auch eine Prozessanalytik Gruppe die sich mit den Daten des gesamten Maschinenparks auseinandersetzt. Je nach Fragestellung werden hier von unterschiedlichen Personengruppen unterschiedliche Teile der Daten analysiert.

S: Habe ich das richtig verstanden es gibt hier noch keinen allgemeinen standardisierten Prozess wie die Daten ausgewertet werden, sondern je nach Fragestellung kommen andere Analyse Prozesse zur Anwendung?

I: Genau je nach Fragestellung muss mit den Daten ganz anders umgegangen werden und daher auch immer neu überlegt werden wie die Daten ausgewertet werden sollen. In der Prozessanalytik wurden viele der Fragestellungen die in Zukunft relevant sein werden noch gar nicht gestellt.

S: Zum Thema Innovation und Optimierung: gibt es hier ein eigenes Projekt, dass sich mit der Fragestellung wie man die Daten für Optimierungsprozesse bzw. Innovation verwenden kann, beschäftigt?

I: Für uns ist die eigentliche Innovation, dass wir uns mit den Daten in dieser Tiefe auseinandersetzen. Wenn ich jetzt aber an die Definition von Optimierung und Innovation die du benutzt zurückdenke, fällt das für mich eher in den Bereich der kontinuierlichen Verbesserung. Da es unser Ziel ist mit den Daten unsere Prozesse weiter zu verbessern. Wir nutzen die Daten in ihrer Gesamtheit um die bestehenden Prozesse zu beurteilen. Das ist für mich die wirkliche Innovation, vor allem da ich auch sehe wie in anderen Bereichen versucht wird, das umzusetzen was im Bereich Datenanalytik vom Marketing versprochen wird ohne grundsätzliche Fragen wie z.b. die Validität der Daten an sich zu klären. Z.b. bei Kurven Daten: diese kann ich als Person sehr gut bewerten bzw. graphische Darstellungen an sich. Das heißt ich kann z.b. ein Bild von dir gut dir zuordnen. Für einen Computer sind das allerdings nur Nullen und Einsen und wenn in diesen Fehlern kommen dann wird es für den Computer schwierig diese korrekt zuzuordnen. Diese grundsätzliche Fragestellung haben meiner Meinung nach viele noch nicht verstanden und verstehen auch nicht wie wichtig diese allerdings in Bezug auf die Beurteilung von Prozessen anhand von Daten sind. Das ist für mich die Innovation in der Arbeitsweise mit Daten.

S: Glaubst du das aus diesen Daten später Innovation entstehen kann. Wenn alle Kinderkrankheiten ausgeremert sind, die Daten alle zur Verfügung stehen. Macht es dann deiner Meinung nach Sinn Big Data Algorithmen einzusetzen um gewisse Zusammenhänge zu erkennen? Und zwar ohne eine konkrete Hypothese die getestet wird, sondern mit einer offenen Fragestellung. Also die Erkenntnisgewinnung aus unstrukturierten Daten?

I: Nein denn hier kommt es zu einem ganz klaren „garbage in - garbage out“ Problem. Man muss konkrete Fragen stellen um konkrete Antworten zu erhalten. Ohne diese Fragestellung denke ich nicht, dass man wirklich verwertbare Informationen gewinnen kann. Das wird zwar oft vom Marketing dieser Software und Datenanalyse Anbieter versprochen, allerdings sehe ich das immer mehr Leute erkennen, dass man 90% der so gewonnenen Erkenntnisse nicht verwerten kann. Von den restlichen 10% sind 9,5 ohnehin bereits bekannt Informationen und vielleicht 0,5% neue Erkenntnisse. Nun kann man natürlich sagen diese 0,5% sind wertvoll und rechtfertigen den Aufwand, allerdings glaube ich, dass mit dieser Herangehensweise zu viele Ressourcen vergeudet werden und man anders bessere Ergebnisse erhält. Vor allem arbeiten wir in einem Bereich der grundsätzlich gut erforscht ist, daher glaube ich nicht, dass eine bahnbrechende Innovation wahrscheinlich ist, oder zumindest nicht, dass man diese so findet. Allerdings ist es für uns ganz entscheidend, dass wir uns mit den Prozessen an sich beschäftigen. Vor allem da nicht sicher ist ob wir in der Zukunft noch genügend genug qualifiziertes Personal finden werden das die Maschinen in dieser Qualität betreiben kann. Hier könnten wir natürlich sagen wir bilden beispielsweise 100 Personen aus die die Prozesse optimal verstehen bzw. versuchen solche Personen zu finden. Diesen müssten wir aufgrund ihrer Begrenztheit natürlich sehr hohe Löhne zahlen. Oder wir gehen den Weg, dass wir versuchen 10 Personen zu finden die diese Prozesse in der von uns gewünschten Qualität verstehen und haben dann noch 90 weitere Personen denen wir es mit Unterstützung durch unser Verständnis des Prozesses ermöglichen mitzuarbeiten. Diese Unterstützung kann dann z.B. durch ein Computersystem bzw. eine künstliche Intelligenz gegeben werden. Das dafür notwendige Wissen basiert auf unserer Projektarbeit jetzt und dem Wissen, dass die zehn hochqualifizierten Personen liefern. Ich denke, dass deshalb die Kenntnisse über den Prozess überlebenswichtig für das Unternehmen sein werden, da anders die notwendigen Arbeitskräfte nicht zur Verfügung stehen werden.

S: Glaubst du das Erkenntnisse aus den Prozess Daten, in welcher Weise auch immer diese gewonnen werden, auch Einflüsse auf andere Bereiche wie z.B. der Konstruktion von Teilen haben werden?

I: Die Datenanalyse ist hierbei nur ein Tool. Aber das ist genau der Weg den wir gehen, wir wollen den Prozess verstehen und zwar in einer Tiefe wie es bisher nicht möglich war. Da befindet man sich dann in eine ganz anderen Welt zBsp im Bereich der Simulationen, In denen analysiert wird, welche Veränderungs-optionen und deren Auswirkungen zur Verfügung stehen. Hier geht es vor allem um digitale Zwillinge. Dabei simuliere ich meinen Prozess allerdings muss ich diesen auch soweit verstehen, dass ich verstehe welche Erkenntnisse ich aus der Simulation erhalte. Z.b. wie valide die Ergebnisse sind. Diesen Fragestellungen haben wir seit 10 bis 15 Jahren, nun kommt es halt unter dem Schlagwort „Digital Twin“. Das ist allerdings unser tägliches Brot mit dem wir immer schon arbeiten.

Zum Thema Produktionsüberwachung: wenn ich den Prozess soweit verstanden habe und auch digitalisiert zur Verfügung habe gäbe es ja auch die Möglichkeit live einen Soll-Ist-Vergleich zwischen dem wie die Prozesse aktuell laufen und wie sie eigentlich laufen sollten durchzuführen. Hierbei könnte man die gesamte Produktionsüberwachung automatisieren. Gibt es hierzu bereits Versuche? Z.b. bei kleineren einfachen Prozessen die bereits sehr stabil laufen.

I: Diese Versuche laufen ständig allerdings ist hier das Problem, dass es so viele unterschiedliche Varianten, wie ein Prozess laufen kann, gibt die vorkommen können. Unsere Herangehensweise ist es zu analysieren welche Fehler Gruppen es gibt und kommen in den Daten diese Fehler Gruppen bei einem spezifischen Prozess vor. Dieser Thematik arbeiten inzwischen sehr große Teams.

S: Gibt es externe Beteiligte an diesen Projekten?

I: Externe Beteiligte gibt es kaum, da dieses spezifische Wissen extern nicht wirklich vorhanden ist.

S: Wie du vorher angemerkt hast ist eure Herangehensweise die Prozesse zu stabilisieren. Wird auch der umgekehrte Weg gegangen in dem Experimente durchgeführt werden bei denen der Prozess bewusst beeinflusst wird?

I: Ja das wird gemacht. Es gibt auch die Fragestellung was für eine Beeinflussung muss durchgeführt werden um einen bestimmten Fehler hervorzurufen.

S: Gab es sonstige Probleme in der Umsetzung dieses Projektes?

I: Natürlich gab es laufend kleinere Probleme unter anderem zum Beispiel die Innovationsfähigkeit unserer Lieferanten diese bauen seit Jahrzehnten Maschinen und für die ist es nicht einfach, sich in ihrer Denkweise, wie so eine Maschine gebaut werden muss, zu ändern. Diese Anlagenbauer sind teilweise gar nicht darauf vorbereitet, dass wir die Daten analysieren, Fehler finden und dann fordern, dass diese Fehler ausgemerzt werden. Ein weiteres Problem war und ist auch die richtigen Personen für dieses Projekt zu finden. Herauszufinden welche Fragen man eigentlich stellen muss. Wobei diese Lerneffekte sich sehr rasch eingestellt haben. Durch die Herangehensweise des Projektteams entstanden viele Lösungsansätze an die vor 2 Jahren noch gar nicht gedacht wurde. Eine wichtige Thematik ist auch die Mitarbeiter die bereits Unternehmen vorhanden sind, mit an Bord zu holen damit diese die Wichtigkeit des Projekts verstehen. Das war am Anfang nicht ganz einfach aber mittlerweile stellen sich hier die ersten Erfolge ein.

S: Das heißt es gibt hier auch erst Erfolge die im Produktivbetrieb verwendet werden?

I: Ja laufend. Obwohl wir in der Sensorik und dem Projekt an sich noch gar nicht dort sind, wo wir in der Zukunft sein wollen, erreichen wir bereits durch die Vorarbeit z.B. stabilisieren der Prozesse, viele Erfolge schon jetzt.

[Interview with interview partner A from Company 2 \(Interview C2\)](#)

Einleitung: Begrüßung und nochmaliges Spezifikation der Unterschiede zwischen Optimierung und Innovation

S: Zuallererst eine sehr grundsätzliche Frage, das Unternehmen stellt Maschinen her, werden in diesen mittlerweile zusätzliche Sensoren verbaut? Beziehungsweise gibt es auch

in den Produktionsmaschinen die die Maschine des Unternehmens herstellen, mittlerweile zusätzliche Sensoren um so mehr Daten zu generieren?

I: Grundsätzlich nicht. Wenn man sich ansieht wie so ein Kunststoffbauteil eigentlich entsteht: Es gibt auf der einen Seite die Spritz Seite, wo das Material über eine Schnecke gefördert wird, auf der anderen Seite das Werkzeug in dem das Bauteil entsteht. Wir kommen sehr stark von der Spritzseite. Hier könnte man z.B. Drücke und Temperaturen über den gesamten Prozess sammeln und daraus Erkenntnisse für die Bauteilqualität ableiten. Das gibt es, allerdings mehr für Versuchs Maschinen z.B. im Haus, aber auch bei Großkunden. Jedoch kaum in der Produktion. Was man aber tendenziell versucht, Sie kennen das vom Reifen, Die Reifendruckkontrolle, hier ist kein direkter Sensor verbaut, sondern indirekt über die Drehzahl des Rades, die bei einem Druckverlust höher ausfallen würde, als die der anderen Räder, wird hier ein Reifendruckverlust festgestellt. Im Grunde versuchen wir es auch so. Wir haben z.B. Sensoren für Schließkräfte, die ohnehin verbaut sind, oder man leitet es aus Drücken ab, die man benötigt um die Maschine zu bewegen. Daraus versucht man dann aus Erfahrungsschätzen bzw. eines aufgestellten Modells, an die eigentlich gewünschten Daten zu kommen z.B. Innendrucke. Hierfür gibt es aber keine eigenen Sensoren. Zusätzliche Sensorik zu verbauen, um an diese Messwerte und auch andere zu kommen, ist eher selten. Wir haben aber natürlich auch viele Sondertechnologien, Spezialanwendungen wo man etwas mehr machen kann.

S: Gibt es eine direkte Verbindung von Maschinendaten und Innovationsprozessen im Unternehmen?

I: Ja wir haben das seit kurzem, allerdings nur testweise bei einigen wenigen Maschinen in Verwendung, dass wir versuchen Maschinendaten direkt vom Feld abzuholen. Um z.B. Lebensdauer Modelle zu entwickeln. Hier wird online über einen Rendezvous-Server oder ähnliches auf die Daten zugegriffen. Diese Online Lösung ist zu jung, davor wurden die Daten von einem Servicetechniker bei der Maschine direkt vor Ort ausgelesen. Wir versuchen hier schon Maschine Daten zu verwenden für z.B. kugelgewindetriebe und ähnliches. Um diese zu verbessern müssen wir natürlich wissen wie diese gefahren werden. Das ist ähnlich wie beim Auto auch hier kann ich im Labor hundertmal messen und auf der Straße sind die

Werte dann ganz andere. Dazu müssen wir wissen wie diese wirklich gefahren werden und das versuchen wir abzuholen. Hier werten wir aus was ohnehin vorhanden ist an der Maschinensteuerung z.b. eingestellte Geschwindigkeiten. Zusätzliche sind zurück gibt es hier kaum, aber es gibt es in Ausnahmefällen z.b. dass man Späne im Öl über einen elektrischen Leitwert misst.

S: Wie werden diese Datensätze analysiert gibt es hier eine Person die mit Hilfe ihres eigenen Wissensstandes versucht anhand der Daten gewisse Verbindungen zu erkennen oder wird das vor allem computergestützten analysiert? Oder sogar mit Big Data Algorithmen?

I: In einem ersten Schritt ist es sicherlich so dass sich die Daten ein erfahrener Mechaniker ansieht. Aber die Arbeiten am Aufbau von solchen Algorithmen und gehen auch viele Kooperationen ein z.b. mit dem Software Center Hagenberg. Zusätzlich bauen wir auch Mitarbeiter auf die schon grundsätzlich von ihrem Interesse in Richtung IT gehen.

S: Das heißt bisher gibt es aber noch keinen strukturierten Prozess wie Daten analysiert werden?

I: Zurzeit noch nicht das ist alles noch sehr tief in der Pipeline drin. Und das Ganze ist doch nicht ganz einfach da man ja schon wissen muss was man eigentlich sehen möchte um mit diesen Algorithmen arbeiten zu können. Da sind wir gerade im Aufbau.

S: Gab es bei der Implementierung irgendwelche Schwierigkeiten, irgendetwas das ganz unvorhergesehen war?

I: Da ich selbst nicht daran beteiligt war kann ich das nicht im Detail beantworten, was ich allerdings aus der Organisationen im Vorfeld weiß, ist das wie so oft in solchen Projekten die Schwierigkeit da war, an die eigentlichen Daten der Maschine beim Kunden zu kommen. Da diese natürlich einwilligen müssen, dass deren Daten verwendet werden dürfen. Umgekehrt gibt es dir dieser Schwierigkeit natürlich auch gegenüber einem Werkzeugmaschinenbauer. Systemseitige Probleme: Grundsätzlich muss ich natürlich auch so sein das eine servicebox

beim Kunden eingebaut wird die von sich aus eine Verbindung herstellt, da natürlich niemand externen Zugriff auf das eigene Netzwerk gewähren möchte. Was auch vollkommen verständlich ist, da ist viel zu gefährlich wäre. Für diese servicebox haben wir einen Entwicklungspartner. Im Allgemeinen ist sicherlich das Kernthema Vertrauen und Mehrwert für alle Beteiligten. Von weiteren technischen Problemen weiß ich nichts. Aber das Projekt ist natürlich ongoing.

S: Gibt es externe Partner außer den bisher genannten z.b. im Innovationsbereich? Ist hier die Tendenz diese Dinge ihr intern im Unternehmen zu behandeln oder sich auch extern Unterstützung zu holen?

I: Um den Prozess zu strukturieren haben wir unterschiedlich Universitäten als Partner. Diese kommen aber vor allem von der technischen Seite. Externe Berater holen wir uns vor allem im Bereich Agile Methoden ins Unternehmen, jedoch nicht unbedingt im Bereich Maschinendaten. Externe Berater in Bezug auf den Innovationsprozess vor allem im Bereich Big Data würde ich eher verneinen. Es gab allerdings vor einiger Zeit ein Projekt wo man sich mit dem Entwicklungsprozess an sich beschäftigt hat, hier waren externe Berater beteiligt. Im Unternehmen werden solche Projekte aber eher intern abgewickelt.

S: Gibt es Experimente, also dass man gezielt versucht bestimmte Parameter zu verändern und deren Auswirkungen auf andere Werte zu messen?

I: Ja das wird sowohl im Unternehmen, als auch dem Kunden durchgeführt. Das bietet auch die Datenbasis. Ja das wird sowohl in den regulären Maschinen Werken durchgeführt als auch in speziellen verschlossenen Entwicklungs-Einrichtungen. Zusätzlich gibt es auch Kundentechniker die z.b. bei Abnahme durch den Kunden die Maschine testen. Auch beim Kunden selbst wird versucht mehr über den Prozess zu lernen und Daten einzuholen.

S: Was sind grundsätzlich die Erwartungen bzw. Ziele vom Unternehmen an Maschinendaten?

I: Ein klares Ziel ist predictive maintenance, dass man mit einer sehr hohen Genauigkeit die Lebensdauer gewisser Teile bzw der Maschine an sich vorhersagen kann. Das ist auch gerade in progress. Bei sehr großen Maschinen die z.b. das Cockpit von Kleinflugzeugen in einem Schuss spritzen, wenn hier die Schnecke gezogen werden muss, steht diese natürlich länger. Hier brauchen Sie einen Kran und auch die richtigen Techniker, das kostet natürlich auch dementsprechend viel. Zusätzlich Ist voraussichtlich auch das Material, dass ich bereits in der Schnecke befindet zerstört. Predictive maintenance würde hier helfen so etwas früh genug vorherzusagen. Zweitens erhoffen wir uns natürlich eine Verbesserung am Prozess selbst. Denn am Ende zählt welches Produkt der Kunde mit unseren Maschinen herstellt und in welcher Qualität er diese fertigen kann. Hierfür haben wir jetzt schon Software-Systeme in der Steuerung der Maschine die sich z.b. damit beschäftigen, dass das Schussgewicht immer gleich ist, dass die Kräfteverteilung immer gleich ist etc.

S: Gibt es bereits Erfolge aus Maschinendaten?

I: Ja das würde ich schon sagen. Auf Messen werden diese Systeme bereits ausgestellt. An der Maschine selbst, wir nennen das IQ Produkte: Intelligent quality. Hier wird zum Beispiel das Schussgewicht gleich gehalten, man beschäftigt sich mit Energieeinsparung. Das nur das benötigte Kühlwasser durch den Kreislauf geschickt wird. Und es gibt auch die ersten Anwendungsfälle von predictive maintenance, hier wird zuerst wie im Maschinenbau das Belastungskollektiv benötigt, das ist ein klassischer Anwendungsfall von Data Mining. Der intelligente Kugelgewindetrieb wird zumindest beworben ob er in Serie ist, bin ich mir nicht sicher. Vom IQ weight control werden rund 1000 Stück pro Jahr mit den Maschinen mitverkauft.

[Interview with interview partner A and B from Company 3 \(Interview C3\)](#)

Einleitung: Begrüßung und nochmaliges Spezifikation der Unterschiede zwischen Optimierung und Innovation

I1: Grundsätzlich vorab: Nachdem es regionale und Geschäftsbereich abhängige Unterschiede in den einzelnen Betrieben gibt. Gibt es im Unternehmen sowohl Betriebe die schon sehr fortschrittlich sind, aber auch solche die mit diesem Thema noch gar keine

Erfahrungen gemacht haben. Gibt es auch keine klassischen Case der für alle zutreffen würde, im Unternehmen.

S: Werden seit dem Digitalisierungsprojekt bzw. schon davor zusätzliche Sensoren in Produktionsanlagen verbaut? Wenn ja wofür, z.b. predictive maintenance?

I1: Ja werden sie, insbesondere in drehenden Bauteilen und solchen mit hohem Verschleiß. Z.b. Kompressoren und Verdichter. Zu den Beispielen kann ich Ihnen zwei bis drei Gruppen sagen. Das erste haben sie schon ganz gut erkannt nämlich predictive maintenance. Wir machen aber auch und wir haben ja Prozessleitsysteme und auch den Automatisierungsgrad haben wir, um die Prozesse zu steuern. Wir machen jetzt noch zusätzliche Sensorik um dort "Advanced process control" aufzusetzen. Das heißt wir können jetzt auch real time optimization anwenden. dafür brauchen wir nicht nur mehr Daten, sondern hier und da auch zusätzliche Sensorik. Also real time optimisation ist unser Schlagwort. Advanced process control & predictive maintenance betreiben wir ebenfalls und für all das nutzen wir auch zusätzliche Sensorik.

S: Diese Sensoren sind im Normalfall ja sehr anfällig für außen Einwirkung und Veränderungen. Gibt es hierbei im operativen Betrieb öfters Probleme. Gehen Sie eventuell auch einen anderen Weg um indirekt Werte berechnen zu können? Also über Werte und Sensoren die ohnehin eingebaut sind und aus Erfahrungswerten und mit Einbezug von mathematischen Modellen die notwendigen fehlenden Werte zu errechnen?

I1: Zu diesem Thema fällt mir ein, dass es im Zentralbereich ein Projekt bzw. einen Sparringspartner gibt, dabei geht es am Ende um künstliche Intelligenz und in Richtung Data Mining. Das machen wir auch allerdings sind wir dort zurzeit noch in den Anfängen. Das ist ein Feld wo wir uns zurzeit noch auf die Werte der Sensoren bzw. das was uns das Prozessleitsystem liefert verlassen. Und weniger auf ein datenbasiertes Model um Daten auszuwerten. Das machen wir auch, da sind wir mit dabei. Allerdings habe ich da jetzt keine richtigen Beispiele mit dabei. Aber wir beschäftigen uns natürlich damit. Im laufenden Betrieb ist das natürlich immer schwierig.

I2: Von der Stabilität der Sensoren her, hätte ich allerdings nicht den Eindruck, dass es hier irgendwo im operativen Betrieb Probleme geben würde. Das sind natürlich aber auch keine Sensoren die man mal eben schnell auf Ebay kauft, sondern hochprofessionell und seit langem bewährte Bauteile. Von spezialisierten Anbietern. Bisher habe ich noch nie das Feedback bekommen, dass wir hier größere Probleme haben würden.

I1: Ich kann hier noch etwas ergänzen: aufgrund der Branche und der Substanzen mit denen wir arbeiten, wenn wir Temperatur, Druck und was auch immer Messen, dann haben wir Schrittketten, Sicherheitsüberprüfungen, jährliche Wartungen etc. Die Sensoren sind auch sehr robust und fallen auch kaum aus.

I2: Ich glaube ich weiß wo sie herkommen mit dieser Frage, bzw. mit diesen Bedenken. Wir haben ja auch immer mal wieder mit Beratern zu tun und die zaubern dann einen neuen Sensor aus der Hosentasche. Allerdings wenn es bei uns um Produktionsabläufe geht, dann reden wir hier um eine ganz andere Art von Sensoren. Ich glaube auch nicht, dass wir hier Anbieter ausprobieren würden, die nicht ohnehin schon aus diesem Bereich kommen.

S: Diese Daten die hier generiert werden, wie werden die weiterverarbeitet? Wie werden sie gespeichert, analysiert und in einem weiteren Schritt dann vielleicht mit einem R&D bzw. Innovationsprozess verbunden?

I2: Ich glaube hier ist das Wording ein bisschen ein Problem, bzw. kommen wir hier eventuell von einer anderen Denkrichtung. Diese Daten gehen natürlich ins Prozessleitsystem ein, aber wir würden hier nicht von R&D sprechen, sondern eher von einer Optimierung. Wenn wir von R&D sprechen, dann ist das für uns eher ein neues Produkt. Einen Prozess mit den Sensoren die in den Prozess eingebaut sind zu optimieren, da würden wir eher nicht von R&D sprechen.

S: Da haben wir eine sehr ähnliche Ansicht, daher meine Frage: wird von der R&D Abteilung auf die im Prozess entstehenden Daten zugegriffen. Um diese dann zBsp für die Entwicklung neuer Produkte zu verwenden?

I2: Das ist eine sehr spannende Frage. Das ist wahrscheinlich eher selten, aber es gibt diese Fälle soviel ich weiß. Es gibt auch eine Person im Unternehmen an die ich gerade denken muss, bei der das in Bezug auf Optimierung, Prozessdaten und Innovation ein bisschen zusammenläuft, da gibt es diese Fälle sehr sicher.

I1: Wir sind auch dort gerade auf dem Weg, diese Daten auf eine Plattform zu bringen. Da probieren wir auch gerade unterschiedlichste Werkzeuge aus. Um dort auf eine Plattform, nutzbar und visualisierbar machen zu können, um sie auch für andere Prozesse und nicht nur für die Prozessoptimierung verwenden zu können. Wir sind dort dabei die richtigen Plattformen uns auszusuchen.

S: Das heißt zurzeit sind die Daten im Prozessleitsystem verfügbar, allerdings gibt es noch keine Möglichkeit, dass Abteilungen und Personen die nichts mit dem eigentlichen Produktionsprozess zu tun haben strukturiert auf diese Daten zugreifen könnten, um diese dann für andere Tätigkeiten zu verwenden.

I1: Vollkommen richtig. Wir haben aber nicht nur das PLS, sondern auch die übergelagerten und das hängt jetzt von Betrieb zu Betrieb ab, immer sehr unterschiedlich. Da haben wir zBsp auch ein MES Layer (Anm: Manufacturing Execution System) darauf. Im gesamten Unternehmen haben wir durchwegs zwar SAP auch noch, allerdings keine durchgängige Landschaft. Wir haben aber also nicht nur das PLS und die Prozessdaten selbst, sondern auch angeknüpft den verschiedenen Layer wie bspw. MES und SAP. Allerdings so wie sie es vorhin in Ihrer Frage formuliert haben, sind wir gerade dabei es auf eine Plattform zu bekommen und für alle nutzbar zu machen.

S: In den Einzelanwendungen bzw. in der Konzeptionierung des Gesamtprozesses: wie wird bzw. werden diese Daten analysiert. Soll es hier ein spezielles Team geben das zBsp in der R&D sitzt oder in der Prozessoptimierung etc. sitzt, oder soll jede Abteilung selbst Datenauswertungen vornehmen und die eigenen Problemstellungen bearbeiten?

I2: Wir sind sehr spezialisiert aufgestellt, daher dürfen sie sich das Unternehmen nicht als zentralistische Einheit vorstellen, wo es dann den oder die Gruppe an zBsp Data Scientists

gibt. Das ist auch eine der Schwierigkeiten für uns, weil man natürlich für viele Sachen dann doch Spezialisten braucht. Für uns ist der Ansatz aber trotzdem weiterhin, dass jede Operationseinheit, jede R&D Einheit sich das selbst analysieren. Vor allem da auch die Technologien zwischen denen sehr unterschiedlich sind. Da gibt es Geschäftsbereiche bei denen ist das schon ein sehr großer Sprung, was die Prozesse und Herstellverfahren betrifft. Da jemanden darüber zu setzen, der versucht beide bzw. auch noch die anderen Geschäftsbereiche zu analysieren, wird kaum möglich sein. Daher wird das eher jeder Geschäftsbereich selbst übernehmen. In den einzelnen Geschäftsbereichen selbst werden zum Start die Fachabteilungen selbst die Analyse vornehmen, wie es dann weitergeht wird sich zeigen.

I1: Wir fangen dezentral an und schauen dann ob es Sinn macht und ob es beherrschbar ist, da eine zentrale Truppe aufzubauen.

S: In der Datenerfassung selbst, werden hier Experiment, also die bewusste Veränderung von Parametern durchgeführt? Im Betrieb oder in Testumgebungen.

I2: In Testumgebungen ja, aber im Betrieb nicht wirklich. Im laufenden Betrieb würde es, wenn dann nur zur Behebung einer Störung durchgeführt, aber wenn der Prozess gut läuft würde man nicht einfach einen Parameter verändern.

I1: Neben dem Angesprochenen, wollen wir aber in der Produktion im Zuge der Digitalisierung, Data Mining und vernetzte Maschinen vorantreiben. Also wir sind auch dort auf der Suche nach "künstlicher Intelligenz" Varianten durchzuspielen und um Abweichungen vom Normalfall zu bewerten. Was das heißt: Im Moment haben wir eine Zeitreihe mit den Daten was passiert ist. Wir wollen natürlich schauen und da geht es in Richtung künstliche Intelligenz und predictive systems, wenn zbsp die Temperatur sich erhöht, dann wird das und das passieren. Um diese Modelle aufzubauen, werden wir Veränderungen, aber eben mehr auf der Datenseite, einzuspielen. Wir haben natürlich auch jetzt schon Simulationstools und Reaktion Simulationen die wir schon jetzt verwenden um zBsp den Prozess vorab zu simulieren.

S: Gab es außer den bisher genannten, Probleme zu denen es im Zuge dieses Digitalisierungsprogramm gekommen ist? Sowohl organisatorisch, als auch technisch.

I2: Jede Menge. Organisatorisch...also die Leute freizuschaukeln, damit sie sich mit dem erst einmal beschäftigen können, ist wahrscheinlich noch um einiges schwieriger, als das vorhin genannte Problem Spezialisten zu finden.

I1: Ja Ressource ist das eine Stichwort, ein anderes Problem war auch, dass wir uns hier in einer sehr heterogenen Datenlandschaft bewegen. Einfach aus der Historie heraus, sind sehr viele unterschiedliche Technologien im Einsatz. wenn wir jetzt probieren, dass zu digitalisieren, müssen wir oft zuerst anfangen, das ganze etwas zu entheterogenisieren und zu standardisieren. Das wird uns auch in Zukunft noch schwerfallen, weil viele Dinge die mal so sind wie sie sind, dann auch für die Zukunft noch "gut" sind, allerdings hindern uns diese in der systematischen Vorgehensweise im Zuge der Digitalisierung. Starke Heterogenität welche Technologien, welche Daten und in welcher Form...es ist einfach ein durcheinander.

S: Gab es externe Beratung und externes Know How das zugekauft wurde? Auch im Rahmen von Projekten mit bspw. Universitäten?

I2: Ja es gab schon eine sehr intensive Phase im Projekt, wo sowohl auf der Markt- als auch auf der Operationseite externe Beratung verwendet wurde. Auf der Operationseite ist das aber nicht nur als Know How zu verstehen, sondern auch Prozessmäßig, also wie baue ich das Programm etc. auf. Wie gestalte ich eine Betriebsanalyse richtig. Das haben wir auch bewusst von dem eigentlichen Projekt getrennt, auch weil das Unternehmen nicht so wahnsinnig viel mit Beratern arbeitet, wenn man das mit anderen Unternehmen vergleicht. In diesem Projekt war die externe Beteiligung für das Unternehmen daher relativ gesehen sehr hoch.

S: Gibt es konkrete Erwartungen und Ziele an das Programm?

I2: Es gibt natürlich ganz viele Einzelprojekte die ein gewisses Ziel, Budget etc. haben innerhalb des Programms, aber das Programm an sich ist so aufgesetzt, dass man gesagt hat:

es ist ein Projekt auf 10 Jahre, da ändern sich die Themen und Schwerpunkte ständig. Und ja es gibt auch ein paar Zielsetzungen was man in 3, 4, 5 Jahren erreicht haben will. Auf der Produktions- und auf der Marktseite, aber das ist eigentlich nicht als hartes Ziel, mit harter Timeline zu verstehen. Sondern mehr als Umbrella.

S: Gibt es schon Erfolge, dass Prozesskosten reduziert wurden, die Qualität erhöht wurde, oder gar ein neues Produkt entwickelt wurde?

I1: Es ist ja so, dass wir das nicht erst seit dem Schlagwort Digitalisierung machen, sondern schon seit Jahren davor. Ich gebe Ihnen hier mal ein ganz konkretes Beispiel, das noch unter dem Begriff Automatisierung, das ältere Wort für Digitalisierung, gelaufen ist. Wir machen schon ewig, dass wir versuchen mit unseren Sensoren, unserem Prozessleitsystem Kosten einsparen. Wir haben jetzt schon hunderttausende, oder Millionen eingespart in dem wir die Daten nutzen und die Prozesse bestmöglich fahren. Jetzt setzt man halt noch eine Ebene drauf, das heißt dann advanced process control, in dem wir die Informationen über einen ganzen Verbund nutzen und damit den Verbund der verschiedenen Anlagen optimal fahren und auch schon vorausschauend regeln und steuern. Also wie in der klassischen Mess- und Regeltechnik. Und hier ist es jetzt wichtig die einzelnen Geräte noch weiter zu vernetzen, noch einen Schritt weiterzugehen. Aber wir sind hier nicht am Anfang und haben auch sehr viel Geld schon gespart, wir wollen aber das Ganze auf den nächsten Level bringen. Noch weiter verbessern. Wir wollen möglichst energiesparend produzieren und hier haben wir noch den Bedarf, dass wir ungeplante Ausfälle reduzieren. Vorausschauend Instandhalten und die Verfügbarkeiten nochmal erhöhen. In dem wir wissen diese Anlage geht bald nicht mehr. Zusätzlich ermöglicht die Digitalisierung auch Arbeitsabläufe zu verändern. So dass wir bei bestimmten Arbeitsabläufen zBsp. Roboter und keine Menschen mehr einsetzen. Es gibt bereits jetzt Anlagen, die sind so weit automatisiert, da ist kein Mensch mehr zu sehen.

I2: Es ist jetzt nicht so, dass wir das Verständnis haben die Digitalisierung hätte vor einem Jahr angefangen und da steigen wir jetzt ein. Es ist aber auch nicht so, dass ist alles schon ein alter Hut und wir brauchen uns darum nicht kümmern. Wir spüren schon, dass sich Dinge vielleicht schneller bewegen, sowohl auf der Markt- als auch Operationseite. Allerdings ist das für uns quasi ein schnelleres Mittel um Kosten zu sparen, die Qualität zu erhöhen. Sprich

die Ziele haben sich nicht großartig verändert, es kommen nur Möglichkeiten bzw. schnellere neue Methoden dazu, um diese zu erreichen.