Importance of Data Quality for Industrial Smart Services

Masterarbeit

zur Erlangung des akademischen Grades „Master of Science (M.Sc.)“ im Studiengang Wirtschaftsingenieur der Fakultät für Elektrotechnik und Informatik, Fakultät für Maschinenbau und der Wirtschaftswissenschaftlichen Fakultät der Leibniz Universität Hannover

vorgelegt von

Name: Schonhoff [redacted] Vorname: Michael [redacted]

Prüfer: Prof. Dr. Michael H. Breitner

Ort, den*: Hannover, den 25.01.2019
*(Datum der Beendigung der Arbeit)

Table of contents

1 Introduction

The issue of industrial smart services is more topical than ever before. The world is developing into an increasingly networked place. With around 15 billion products connected via the internet in 2015, it is estimated that it will reach approximately 30 billion in 2020. Digitization is thus one of the biggest topics that is currently haunting the economy. This is accompanied by new future projects such as “Industry 4.0”.
Many new technologies and approaches are causing disruption – including smart services. In times of increasing globalization and the associated tougher competition, smart services offer companies great opportunities. For instance, on the user side tailor-made, intelligent services with increased added value, which can lead to higher productivity. At the same time, the provider collects data about the user and, through suitable analysis of the data, can offer even better, adapted services in the future. Thereby, the customer loyalty is further increasing.
However, there are many different challenges to be mastered. One of them is the acquisition and use of high quality data as data are fundamental for efficient and well functioning smart services. While there are many scientific works that examine the topic of data quality, the dimensions of data quality and its importance in the context of industrial Smart Services has not been researched yet.

1.1 Problem and Relevance of the Thesis

Considering the fact that competition between companies is becoming ever tougher, smart services offer a way to gain a competitive edge over competitors. Therefore, the further investigation of smart services is important. Data quality plays an essential role as a basic building block of smart services. That is why, in this thesis, the dimensions of data quality that are relevant for industrial smart services are identified in order to make data quality tangible. In a second step, the dependencies between the dimensions are described and the importance of each dimension is derived from a use case example. On the basis of these findings, a recommendation of action in form of a process is worked out that helps companies to develop smart services.

1.2 Motivation and Objective

Identifying dimensions of data quality in the context of industrial smart services helps companies develop and implement these. Smart services offer the opportunity to gain information about the customer and thus to better understand the customer. This enables the company to understand the customer and his needs better and thus satisfy them with tailor-made services. It follows that customer value increases, which leads to greater customer satisfaction and thus to greater customer loyalty. This allows companies to generate a competitive advantage in the competitive market.
Since the mid-nineties, the issue of data quality in the scientific field has become an ever-widening field of interest. The requirements of data quality depend on the field of application. In the context of industrial smart services, this topic has not yet been addressed. However, data quality is the foundation of custom services. Thus, answering the questions

1) What are the dimensions of data quality in the context of industrial smart services?

   and

2) What significance does data quality have for industrial smart services?

can close an important research gap. Based on the dimensions identified, companies can make data quality quantifiable. Furthermore, the dependencies between the dimensions are illuminated. Based on the derivation of the importance of data quality for industrial smart services in a use case, a recommendation for action is derived as to how companies consider data quality as a decisive factor for smart services already during the development of the services and thus ensure its later application. Thus, the work serves as a starting point for further research in the context of industrial smart services, which are explained in chapter 8.

1.3 Structure of the Thesis

The theoretical background is explained in chapter 2. First of all, the changes and challenges posed by digitization for companies are described. Subsequently, smart services are described as a possibility of a competitive advantage, highlighting the relevance of the topic. Furthermore, the functionality of smart services and the needed infrastructure is described. Subsequently, a structured literature analysis according to Webster and Watson (2002) identifies all relevant dimensions of data quality mentioned in science. Various filters are used to extract the dimensions relevant for industrial smart services from all of the above. Then the dimensions are defined more closely and the dependencies among them are described. Furthermore, the identified dimensions are examined on the basis of a use case. It will be deduced which dimension has which concrete influence in the practical application of industrial Smart Services. Finally, the findings will be used to develop a recommendation for action in the form of a process. This describes how the aspect of data quality can be systematically taken into account during the development of smart services, so that the data quality is ensured during operation. Finally, the limitations and the critical appraisal of the work are executed. Finally, Chapter 8, "Conclusion and outlook," summarizes the results and research questions that build on the findings.

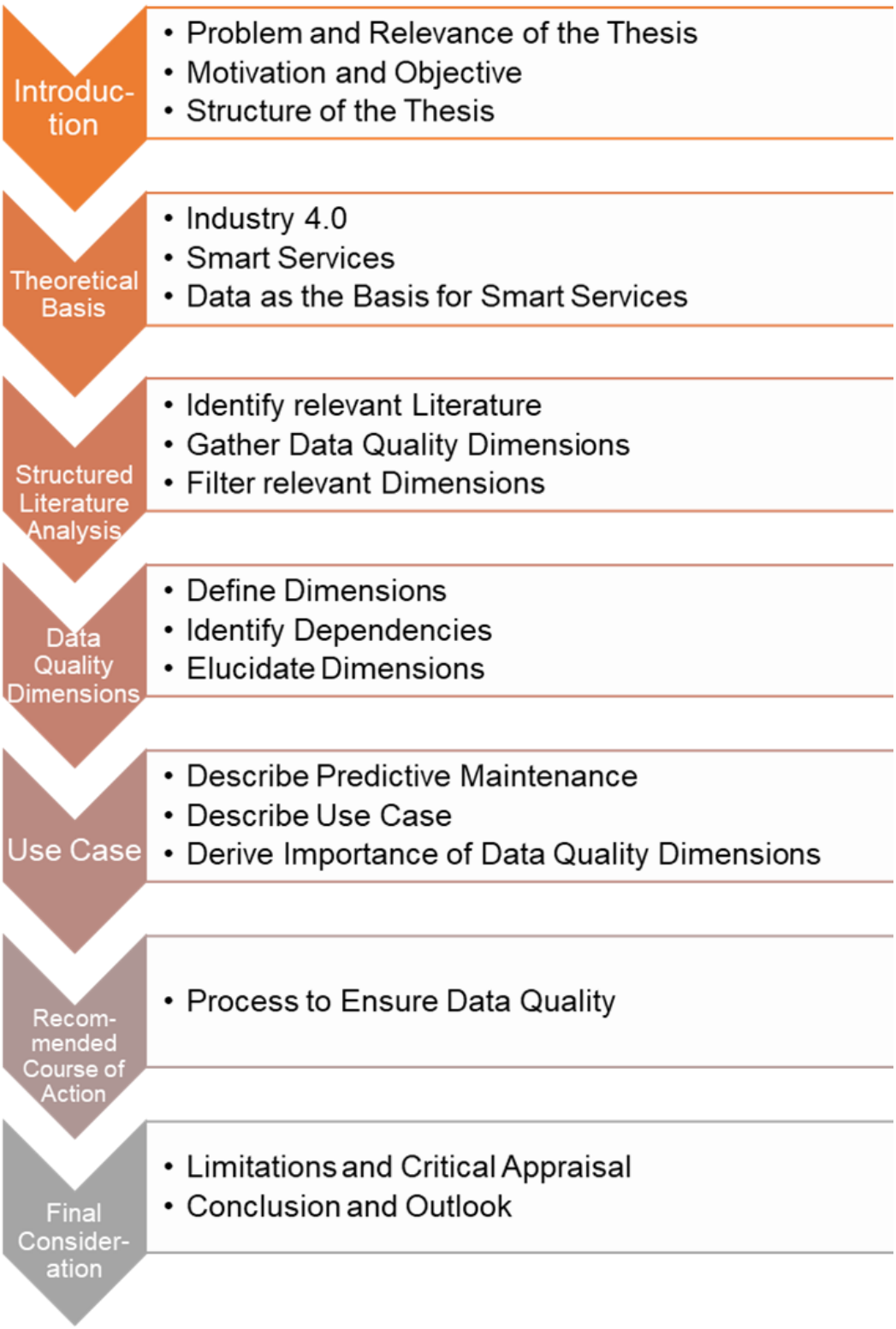

Fig. 1: Structure of the thesis

Source: Own representation

8 Conclusion and Outlook

The digitization of the world is progressing, making it an increasingly interconnected place. Whilst the opportunities are being exploited intensively in some areas, the maturity of digital business models in the industrial sector is still relatively weak. The survival of many companies depends on their adaptability to the digital world. One way for businesses to benefit through digital business models is the smart services. Therefore, a more in-depth consideration of these makes sense.
The basis of smart services is data. Digitization makes it possible to skim huge amounts of data. It is important for smart services to be able to efficiently extract the right information from the data. This requires a high data quality. Depending on the field of application and the viewing angle, data quality is characterized by different properties.
This work focuses on this point. In the literature, data quality has been considered from many angles, but not yet in the context of industrial smart services. In order to identify the relevant dimensions, a structured literature analysis according to Webster and Watson (2002) was carried out. First of all, all the dimensions mentioned above were determined, and filters were finally used to identify the dimensions relevant to industrial smart services, namely accuracy, completeness, timeliness, relevance, accessibility, security, ease of manipulation and credibility. The identification of these dimensions helps in practice to bring the term "data quality" to life and to make it tangible by enabling the dimensions to clearly define and quantify individual aspects of data quality. For example, when developing smart services based on dimensions, you can develop requirements that can be systematically addressed. The concrete requirement for a certain dimension and the weighting of its importance always depend on the individual case of application.

It was developed that the dimensions cannot be considered in isolation from each other, because they are in their respective dependencies. These dependencies were also worked out here. Understanding dependencies in both the development of smart services and error analysis helps when data quality issues arise. Thus, a sensory measuring system could technically meet the requirements of the dimension accuracy, but e.g. failure to consider all the relevant influencing variables does not provide valid data. During development, the identified dependencies are used to determine which dimensions must be taken into account when achieving a dimension. In the case of error analysis, the dependencies help to find the root of an occurring error. Furthermore, the importance of data quality was analyzed by means of a theoretical

example. The result shows that satisfactory data quality is essential for working industrial smart services. Thus, it has been shown that any identified relevant dimension can potentially contribute to severely restricting a service's performance or even stopping it if the dimension has not been adequately met. For example, the dimensions accuracy or completeness directly affect the information capability of the data, while dimensions such as accessibility or ease of manipulation tend to focus on the processability of the data. Both categories are essential for good smart services. Although the work is subject to some limitations due to its theoretical character, a recommendation for action was finally developed in the form of an illustrated process flow. This process is addressed to decision-makers in the field of industrial smart services development such as predictive maintenance. This process provides the addressees with orientation in the development of smart services with regard to the consideration of data quality. This process describes in step-by-step ways how data quality during development can be sufficiently taken into account to ensure the quality of the data in smart service operation.

Starting from the limitations mentioned in chapter 7, there are some challenges for future research. Since this work has examined the dimensions of data quality in the context of industrial smart services, it would be interesting to examine whether the theoretical knowledge gained works in practice. On the basis of these findings, it would then be possible to investigate whether companies should be considered in terms of their properties, such as size, industry sector, etc., in order to derive the requirements for data quality dimensions. This would allow a framework to be developed that would make it easier for companies to identify their requirements for the different dimensions. Because cooperation between companies is a central aspect of smart services, standards for dimensioning would make sense. This would allow a set of best practices and KPIs to be defined in order to achieve the respective data quality dimensions or measure their performance.

Summing up, this thesis is a good initial point for continuative research. The findings and concepts developed in the work serve as the basis for the development of high-performance industrial smart services and can thus be seen as a valuable contribution to industrial companies who want to make the transition from a mere manufacturing company to a service provider.