Tampere University

Sofi Sulanen

IMPROVING FINANCIAL DATA QUALITY THROUGH DATA GOVERNANCE

Faculty of Management and Business
Master's Thesis
April 2021

ABSTRACT

Sofi Sulanen:	Improving financial data quality through data governance
Master's Thesis	86 pages, 1 appendix
Tampere University	Master's Degree Program in Business Studies, Accounting
Supervisor:	Professor Lili Kihn
Published:	April 2021

Organizations around the world aim to become data-driven and derive competitive advantage of business data to succeed in the challenging environment. Data is viewed as an important resource and an asset in companies but the quality of data is often not paid enough attention to. In reality, organizations are often unaware of the quality of their data (O'Brien 2015, 443). The quality of financial data is especially important for companies because it is used in business decision making and external reporting. However, financial data in companies is rarely governed in the same way as other business assets.

The purpose of this research was to study how the quality of financial data can be improved by utilizing data governance to address data quality challenges. The theoretical framework is composed of data quality and data governance literature. The themes are first clarified separately and thereafter, a synthesis is made of them in the summary of the theoretical framework. The research was conducted as a qualitative case study in which a design-based research approach was used. The empirical data was collected from eight semi-structured interviews, where the case company's employees were interviewed. The interviewees were selected from financial accounting and management accounting teams to represent the main stakeholders of financial data. The interview data was used to gain understanding of the data quality challenges and the current state of data governance in the case company. In addition, the interviewees were asked to consider their needs regarding the governance of financial data and the benefits they expect to gain from better governed data.

From the interview data, five main themes of challenges in the current state of the company were identified. The themes included management, roles and responsibilities, communication, internal conditions and technology related challenges. Due to these challenges, the requirements for data quality and the employees' responsibilities regarding data were not clear which negatively affected the quality of financial data. In addition, decision-making authority had not been defined in the company which created a risk for data quality if several people were making decisions of the data individually. Because the identified challenges were mostly organizational instead of technical, data governance was seen as a suitable solution to address the challenges. Based on the identified challenges and the needs from the interviews, a data governance framework was developed for the case company. First, roles and responsibilities regarding financial data were defined. Then, data governance activities were designed to document the common principles for working with financial data and to enhance common understanding among data stakeholders.

The findings of this research imply that data governance can be used to improve the quality of financial data in organizations because it addresses the organizational challenges that negatively affect financial data quality. The research was restricted to studying a single company and the data governance framework was developed explicitly for the case company. Therefore, the findings of this research cannot be generalized to other organizations. However, this research contributes to the literature by increasing understanding of the challenges for financial data quality and utilizing data governance in the context of financial data. In addition, practitioners can use this research as a case example for designing their own data governance activities for financial data. For the case company, developing the data governance framework was the first step towards ensuring high quality of financial data. However, its effectiveness still highly depends on how well it is implemented and adopted.

Keywords: Data quality, Data governance, Data management, Financial data, Design-based research

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

TIIVISTELMÄ

Sofi Sulanen: Talousdatan laadun kehittäminen datan hallinnointimallin avulla
Pro gradu -tutkielma: 86 sivua, 1 liitesivu
Tampereen Yliopisto Kauppatieteiden tutkinto-ohjelma, yrityksen laskentatoimi
Ohjaaja: Professori Lili Kihn
Julkaisuaika: Huhtikuu 2021

Organisaatiot ympäri maailmaa pyrkivät toimimaan dataohjautuvasti sekä tavoittelevat kilpailuetua liiketoimintadatasta pärjätäkseen haastavassa kilpailuympäristössä. Data nähdään yrityksissä keskeisenä resurssina ja omaisuutena, mutta datan laatuun ei usein kiinnitetä riittävästi huomiota. Todellisuudessa organisaatiot eivät usein ole tietoisia datansa laadusta (O'Brien 2015, 443). Erityisesti talousdatan laadun merkitys yrityksille on tärkeä, koska sitä hyödynnetään sekä päätöksenteossa että ulkoisessa raportoinnissa. Kuitenkaan talousdataa ei usein hallita samalla tavalla kuin muuta yrityksen omaisuutta.

Tämän tutkimuksen tarkoituksena oli tutkia, miten talousdatan laatua voidaan parantaa hyödyntämällä datan hallinnointia vastaamaan datan laadun haasteisiin. Tutkielman teoreettinen viitekehys koostettiin datan ja talousdatan laadun sekä datan hallinnoinnin kirjallisuudesta. Teemat esiteltiin ensin erikseen, jonka jälkeen niitä käsiteltiin yhdessä teoreettisen viitekehyksen yhteenvedossa. Tutkimus toteutettiin laadullisena tapaustutkimuksena, jossa lähestymistapana käytettiin kehittämistutkimusta. Empiirinen aineisto kerättiin kahdeksasta puolistrukturoidusta haastattelusta, joissa haastateltiin case-yrityksen työntekijöitä. Haastateltavat valikoitiin ulkoisen laskentatoimen sekä sisäisen laskentatoimen tiimeistä, sillä he edustivat keskeisimpiä talousdatan sidosryhmiä. Empiiristä aineistoa käytettiin lisäämään ymmärrystä talousdatan laadun haasteista sekä datan hallinnoinnin nykytilasta case-yrityksessä. Lisäksi haastateltavia pyydettiin kertomaan heidän toiveensa datan hallinnointiin liittyen sekä pohtimaan etuja, joita he odottavat paremman datan hallinnoinnin myötä.

Haastatteluaineiston perusteella tunnistetut haasteet datan laadussa jaettiin viiteen teemaan, jotka olivat johtamiseen, rooleihin ja vastuisiin, kommunikaatioon, teknologiaan sekä sisäisiin olosuhteisiin liittyvät haasteet. Havaittujen haasteiden vuoksi datan laadun vaatimukset sekä työntekijöiden vastuut talousdataan liittyen eivät olleet selvät, mikä vaikutti negatiivisesti datan laatuun. Lisäksi päätöksentekovaltuuksia ei ollut määritelty, mistä syntyi riski talousdatan laadulle, jos useat ihmiset tekivät erillisiä päätöksiä datasta. Koska tunnistetut haasteet olivat pääasiassa organisatorisia eivätkä teknisiä, datan hallinnointi nähtiin soveltuvana ratkaisuna haasteisiin vastaamiseksi. Tunnistettuihin haasteisiin sekä haastateltavien toiveisiin perustuen yritykselle laadittiin datan hallinnointimalli. Ensimmäiseksi määriteltiin roolit ja vastuut talousdataan liittyen. Toiseksi datan hallinnoinnin liittyvät tehtävät suunniteltiin dokumentoimaan yhteiset toimintatavat talousdatan kanssa toimimiseen sekä parantamaan yhteistä ymmärrystä talousdatasta organisaatiossa.

Tutkimuksen tulokset osoittavat, että datan hallinnointia voidaan hyödyntää talousdatan laadun kehittämisessä, sillä se vastaa niihin organisatorisiin haasteisiin, jotka heikentävät talousdatan laatua. Tutkimus rajoittui tarkastelemaan yhtä yritystä ja datan hallinnointimalli kehitettiin kyseisen yrityksen tarpeisiin. Siten tutkimuksen tulokset eivät ole yleistettävissä muihin organisaatioihin. Tutkimuksen kuitenkin voidaan nähdä lisäävän ymmärrystä talousdatan laatuun vaikuttavista haasteista sekä datan hallinnointimallin käytöstä talousdatan kontekstissa. Lisäksi yritykset voivat hyödyntää tutkimusta esimerkkinä talousdatan hallinnoinnin suunnittelussa. Case-yritykselle datan hallinnointimalli on askel kohti talousdatan laadun varmistamista, mutta sen vaikuttavuus riippuu siitä, miten se implementoidaan käytäntöön.

Avainsanat: Datan hallinnointi, Datan hallinta, Datan laatu, Talousdata, Kehittämistutkimus

TABLE OF CONTENTS

1 INTRODUCTION

1.1 Importance of the topic

This study focuses on examining data governance in the context of financial data. Data has become an essential resource for companies and it is often in a critical role for companies' success. Data are facts of real-life objects that can be turned to information by processing data in a meaningful way. (Atkinson & McGaughey 2006, 85–86). Because making fact-based decisions is a prerequisite for succeeding in the competitive business environment, companies must be able to utilize their data appropriately. Therefore, data is viewed as the main source of sustainable competitive advantage for companies. Luckily, nowadays a significant amount of data is collected of organizations' operations and stored in corporate data centers. In addition, the amount of said data grows in a fast pace, so the lack of data should not be a problem. (Tallon, Ramirez & Short 2013, 142.)

Despite the growing amount of data, it seems that companies do not pay enough attention to the quality of their data. As data is utilized in nearly all operations in companies, it is clear that the data must be accurate and reliable. Hence, high quality data is a prerequisite for effective decision making and operations. In contrast, poor quality data can have a significant negative effect on companies' performance. When poor quality data interferes with the daily operations of a company it negatively impacts decision making, company performance, customer satisfaction and employee engagement. (Haug, Zachariassen & van Liempd 2011, 169–173.) Relying on poor quality data in companies, such as banks may also have significant economic and social effects (Wang & Strong 1996, 6). However, the researchers agree that the importance of data quality has not been fully understood in companies (Haug et al. 2011, 170). In fact, companies are often unaware of the quality of their data and the problems it might cause (O'Brien 2015, 443).

Financial data is especially critical to companies because it is used for external reporting as well as internal support for decision making. For example, financial statements are an important source of information to investors and other stakeholders and thus, the information must be reliable. (Du & Zhou 2012, 76.) In addition, the decisions that are made based on data are only as good as the underlying data (Cheong & Chang 2007, 1000). This emphasizes the importance of paying attention to the quality of financial data because it is clear that the quality of the data in the accounting systems affects both external compliance and internal business support (Bai, Nunez & Kalagnanam 2012, 453). Financial data is often stored and used in information systems, such as enterprise resource planning (ERP) systems which combine company data into one single system and thus, can provide real-time data for different uses (Malinić & Todorović 2012, 723–724). However, the challenge that emerges is how the quality of financial data can be ensured. When a large amount of data is stored in one system, the information is readily available for all the users, but so are the defects in data (Granlund & Malmi, 2002, 304).

Therefore, it can be agreed that the large amount of data alone does not satisfy the information needs required for decision making but rather that the data is of high quality. This requires that the data is well managed and clear principles among data users exist for the users to be able to rely on the data. Hence, the growing amount of data has made data governance an important issue in companies (Tallon et al. 2013, 142).

Data governance is a framework that guides people's actions regarding data to ensure that organization's data assets are properly managed (Ladley 2012, 11). Data assets refer to the data that a company owns and seeks to benefit from (Atkinson & McGaughey 2006, 88). Data governance clarifies who is authorized to make decisions on certain data and what kind of activities result from these decisions. The purpose of data governance is to maximize the value of a company's data assets by making sure that all data activities are conducted in a way that supports high data quality. (Otto 2011a, 241.)

Although the research on data governance has grown during the last 10 years it remains as a rather novel research area (Alhassan, Sammon & Daly 2019, 98). Also, Weber, Otto and Österle (2009, 6) note that compared with IT governance the research on data governance is still in its early stage. This is because managing information used to focus on managing IT resources (Tallon et al. 2013, 149). The researchers agree that nowadays data governance is gaining importance in organizations (Abraham, Schneider & vom Brocke 2019, 424). However, according to Alhassan et al. (2019, 98) data governance should still be paid more attention to in the academic as well as practitioners' community.

The importance of data governance can be seen from the number of regulations passed in the 21st century for governing data which include the Sarbanes-Oxley Act, Basel III and GDPR. The Sarbanes-Oxley Act was passed in 2002 in the US to protect investors after the financial scandals in the US (Schreider 2020, 178). It requires that the reliability of financial reports and the information systems of public companies are personally certified by the CEO and CFO (Bai et al. 2012, 453). In addition, the Sarbanes-Oxley Act requires that IT systems are managed in a transparent manner and internal controls are established to prevent misuse and fraud (Cheong & Chang 2007, 1000). Basel III was created in 2010 for the financial sector and it requires that financial institutions evaluate their technical risks, establish controls and conduct audits to mitigate risks from fraudulent activities and system failures. In 2016, GDPR was approved to require that organizations manage personal data of EU citizens in a way that protects their privacy. (Schreider 2020, 102, 160–161.)

Practitioners have also recognized the value of data governance. In their study Khatri and Brown (2010, 148) refer to a survey conducted 2006 in North America in which companies using business intelligence systems named data governance as one of the most important success factors for leveraging the value of their data. In addition, professional firms, such as KPMG and TietoEVRY highlight the importance of data governance for companies that seek to benefit from their data (KPMG 2020; Aula 2020). In a blog by TietoEVRY, Aula (2020) states that for a company to be able to utilize its data and become a data-driven organization, data governance must be in place. KPMG's report from 2020 points out that as businesses become more and more digital, the traditional governance models are not enough to cover the need for new management and control practices. In addition, according to the report, the responsibility of a company's data is often placed on IT and the CIO. (KPMG 2020, 22, 26.) However, as business functions are often the main creators and users of data, the responsibility for data should be placed on the relevant functions. Regarding financial data, the finance department and the CFO are accountable for the accuracy of financial reporting and thus, they should also be responsible for financial data. (Cheong & Chang 2007, 1001.) Data governance is needed to clarify responsibility areas for data in companies.

Benfeldt Nielsen (2017, 131) found that only few studies in data governance domain use practice-oriented methods in research, such as action research or design science. She states that there is a need for more studies of this kind because of the potential benefits that can be leveraged from implementing data governance in practice. Additionally, most of the research is conducted from the perspective of information systems and computer science. (Benfeldt Nielsen 2017, 131.) Even though researchers agree on the importance of ensuring high quality in financial data and the usefulness of data governance for achieving that, only limited amount of literature discusses data governance as a method to improve and ensure data quality in the context of financial data. To contribute to this gap in the literature, this research focuses on improving the quality of financial data with data governance framework from the perspective of accounting.

1.2 Research objective and restrictions

The objective of this research is to examine how the data governance framework can be utilized in the context of financial data. The purpose of data governance is to affect the root causes of data quality problems and therefore, this research seeks to find out how data quality problems in financial data can be addressed with data governance framework. To answer this question, it is relevant to understand what kind of challenges exist regarding the quality of financial data and how those challenges can be addressed with data governance framework. The main research question that this paper seeks to answer is the following:

How can the data governance framework be utilized to address challenges in financial data quality?

The main research question is divided into two sub questions:
1) What kind of challenges regarding the quality of financial data exist?
2) How can the challenges in financial data quality be addressed with data governance framework?

Whereas most of the previous research on data governance is concerned data in general, this research focuses on financial data. The scope of the data covered in this research is the financial data in a company’s ERP system. Financial data is defined in this research as the data that directly affects the financial processes of a company. This includes master data, such as customer data and transaction data, that is generated in business transactions, such as invoices. Product data and production data have been excluded from the examination in this research because product data does not have a direct effect on financial processes.

The emphasis of data governance is on the organizational processes around data, but it also includes the necessary consideration of the physical IT artifacts, such as the software and hardware that the data is stored in. For example, access rights to certain data are often managed through the software which makes it an integral part of data governance. (Tallon et al. 2013, 149.) Therefore, the focus of this study is on the organizational factors influencing financial data quality. The necessary technical factors and information system aspect are taken into consideration, but this research does not seek to address the challenges by developing the existing IT infrastructure.

This research covers both financial accounting and management accounting contexts in the accounting research domain. That is because both of them are essential uses of financial data and thus neither can be excluded from the research. Financial accounting and management accounting are also closely interrelated from the perspective of data since even though the reports from each generate information for different uses, they might often use the same data. Thus, financial data quality inevitably influences both financial as well as management accounting.

The research is conducted as a commissioned research and therefore, the examination is restricted to a single company. The case company is a Finnish manufacturing company, Framery Oy. The company operates in the furniture industry and manufactures soundproof phone booths and private spaces for open offices. (Framery 2020.) Framery Oy is a suitable case company for this research because it has not developed governance practices for financial data and has observed problems arising from that. The case company is interested in implementing practices that ensure high quality of financial data in order to make better decisions based on reliable information. The company believes that the future of a finance function is a holistic business support in decision making as well as compliance. Thus, Framery sees that it is essential that data competencies are built for people working with financial data. In this research, a data governance framework will be developed for the case organization in order to create structure in data processes and facilitate common understanding of financial data in the organization. This in turn is expected to improve the quality of financial data as well as found a base for the finance function to further grow their data capabilities.

1.3 Research methodology

This research is conducted as a qualitative case study. Qualitative research aims to a holistic examination of the researched phenomenon. The approach of a qualitative study is examining the real world and it seeks to provide in-depth understanding of the researched phenomenon. (Hirsjärvi, Remes & Sajavaara 2009, 157.) The methodological approach in this study is based on hermeneutics. Hermeneutic approach refers to interpretation of human action as a method to gain understanding of the researched phenomenon. It focuses on people's communication and emphasizes subjective interpretations that individuals make of their social world. (Bryman & Bell 2015, 28–31). The chosen methodology supports the purpose of this research because data governance considers the human aspect of data and hence, the phenomenon can be better understood by examining people's perceptions. To address data quality problems with data governance, understanding of the problems must be acquired. Therefore, a qualitative case study provides the needed in-depth information for gaining the necessary understanding.

The research is conducted as a commissioned research for a company where the goal is to improve the quality of financial data. Therefore, this study takes the approach of a design-based research. The goal of a design-based research is to produce a well-argued solution designed particularly for the problem of the research subject. Thus, the created solution is company-specific and the same solution may not work in any other organization. (Tamminen 1993, 157.) Therefore, a design-based research does not seek for generalization of the findings but can still contribute to the understanding of the researched phenomenon.

Research methodologies in the Finnish accounting research are traditionally categorized following the classification by Neilimo and Näsi (1980). Their classification includes four different approaches: nomothetic, decision-oriented, action-oriented and conceptual approach (Neilimo & Näsi, 1980). Later, Kasanen, Lukka and Siitonen (1991) added constructive approach as a fifth approach to the classification.

The design-based research approach used in this research can be seen to have characteristics of an action-oriented and a constructive approach. Action-oriented approach aims to obtain in-depth understanding of the researched phenomenon by examining research subjects on a detailed level. In studies that follow action-oriented approach, empirical data is often collected from one or a small group of research subjects and thus, case studies are commonly used. (Kihn & Näsi 2010, 47–49.) The common characteristic with action-oriented and design-based research approaches is collecting in-depth information about the researched phenomenon possibly in the form of a case study. However, in a design-based research this information is primarily used to gain understanding of the company's problem (Tamminen 1993, 159), whereas in action-oriented research this information is used to gain understanding of the researched phenomenon. The purpose of a study following a constructive approach is to provide theoretical contribution by innovating novel constructions (Kihn & Näsi 2010, 48). Even though the goal of a design-based approach is creating a solution for the case company's problem, the design-based approach differs from the constructive approach because research-based research does not aim for a solution that could be generalized to other companies, which is the aim of the constructive approach (Tamminen 1993, 157).

Case study was chosen as the research strategy for this study. Case study seeks to achieve detailed information of one or a small group of cases that are chosen for the study. Case study aims to gain rich understanding of the researched phenomenon by using empirical evidence that is collected in natural settings. The empirical evidence is often collected by several methods, such as interviewing and observing. (Hirsjärvi et al. 2009, 130–131.) Case study fits as a research strategy for design-based research because it enables examining the case company on a detailed level to improve understanding of its situation and to develop a solution for its problem.

1.4 Structure of the research paper

The research paper is structured as follows. In the first chapter, the topic and the motivation of the study are introduced. Then, the research problem and the research questions which further guide the study are presented. The choices for the research methodology are then presented and the structure of the research is clarified.

The second chapter covers the theoretical framework of the study and consists of two themes, data quality and data governance. Here, the key concepts of financial data quality and data governance are clarified in order to give an overview of the literature that the research is based on. Thereafter, a synthesis of is made of these in the summary.

The third chapter clarifies the choices that were made in the empirical research. First, the case company is introduced to provide understanding of context of the research. Second, the design-based research approach is introduced and the research process is described. Third, the chosen data collection and analysis methods are clarified and then, the reliability and validity of the research are evaluated.

The fourth chapter consists of the findings of the empirical research. First, the empirical data set which consists of the interviews is described. Then, the findings are presented with the aim to answer the first research question from the empirical point of view. Drawing on the findings, a suggestion of the data governance framework for managing financial data in the case company is presented.

The fifth chapter includes discussion and conclusions of the research where the findings presented are discussed and reflected on the theoretical framework. Finally, the implications, limitations of the research and possibilities for future research are discussed. The structure of the research paper is presented in figure 1.

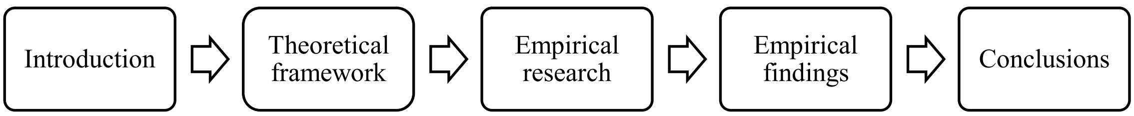

Figure 1. Structure of the research paper

2 THEORETICAL FRAMEWORK

This chapter forms the theoretical framework for the research and consists of two parts. First, the concept of data quality is discussed to provide an understanding of what data quality consists of in the context of financial data. Second, the concept of data governance is introduced because it forms the foundation for achieving high quality data. Also, two models for designing data governance framework are introduced. Finally, a synthesis of data quality and data governance is made in the summary of the theoretical framework

2.1 Data Quality

2.1.1 Definitions for data, quality, and data quality

In this section, the definitions for data, quality and data quality are discussed and different perspectives to data quality are introduced. This provides a basis for understanding data quality in different contexts.

Data can be defined as "facts represented as text, numbers, graphics, images, sound or video" (Mosley 2008). The purpose of data is to represent the characteristics of objects and events by defined symbols. Data can be viewed as the raw material of information because when data is processed in an appropriate way, it can be turned to information. (Ackoff 1999, 170.) Due to the nature of data, same data can be used several times and for different purposes (Tayi & Ballou 1998, 56).

Data can be classified to different categories according to its type. According to Vayghan, Garfinkle, Walenta, Healy & Valentin (2007, 671) the data in companies is commonly classified to master data, transactional data and historical data. Haug, Arlbjørn, and Pedersen (2009, 1055) argue that historical data is a form which the two former data categories can obtain and that it should not be regarded as a category on its own. Thus, they only recognize master data and transactional data in their study. (Haug et al. 2009, 1055.) Master data refers to the company's key external and internal objects, such as customer, vendor and product data (Vayghan et al. 2007, 671). Transactional data represents information of business transactions, such as sales orders and invoices. Transactional data always includes a time dimension, a value and references to other data (Haug et al. 2009, 1055). References to other data can include master data objects, such as customers, to which sales invoices are linked to. These references can also include master data dimensions, which can be used to categorize other data, such as an account. Some researchers use the term reference data for this data type and regard it as a subset of master data (Allen & Cervo 2015, 19; Dreibelbis 2008). Thus, master data creates the base data to which transaction data is connected to (Haug et al. 2011, 169). Dreibelbis (2008) notes that defining the data that is classified as master data is not universal but differs among industries and companies.

Today, data has become an essential resource for companies (Atkinson & McGaughey 2006, 85). Data is collected and used in numerous business activities and decision making is based on data (Haug et al. 2011, 169). Thus, data should be treated as an important company asset instead of a byproduct residing in information systems (Ladley 2012, 9–11). Asset in this context does not require a physical form but refers to the economic value that data holds and can be derived from it. Thus, data assets refer to the integral information resource that companies have and are able to financially benefit from when data is utilized in an appropriate way. It has also been suggested that data should be included in balance sheet among other company assets. (Atkinson & McGaughey 2006, 88, 94.)

The idea of perceiving data as an asset emphasizes the importance of the quality of data. The higher the quality of the data is, the more valuable data assets are because data can be used effectively. In contrast, poor quality data diminishes the value of organization's data assets (Even & Shankaranarayanan 2007, 80–81). Therefore, in order to improve data quality, it is important to understand what it consists of.

There is not a single common definition for the general concept of quality because quality can mean different things depending on the perspective it is viewed from. Kihn (2015) studied the concept of quality from three perspectives which included technical quality, commercial quality and service quality. Historically, quality was viewed narrowly from the technical quality perspective which referred to internal flawlessness of the object. Later, the perspective was broadened to also include the ability of the object to fulfill the user's requirements. (Kihn 2015, 291–292.)

These narrow and broad perspectives to quality are also discussed in the data quality literature. Traditionally, the research on data quality focused on objective perspective to quality, which refers to viewing data quality independent of the context and assessing it by objective quality attributes, such as accuracy (Strong, Lee & Wang 1997, 104). Wang and Strong (1996, 6–7) argued that without consideration of context of the data use, the research on improving data quality focused on developing the objective characteristics of data quality which did not provide real value. They suggested that improving data quality requires understanding the data user's requirements to data and assessing data quality from the user's perspective. Thus, they adopted the concept of "fitness for use" from quality literature to data quality. (Wang & Strong 1996, 6–7.) Since then, assessing data quality as fitness for use has been widely accepted as the classic perspective for assessing data quality (Haug & Arlbjørn 2010, 292).

Even and Shankaranarayanan (2007, 78) elaborate the understanding of the user's perspective to data quality by describing factors that influence the context where data is used. First, scope refers to the different assessments that an individual, a department and a company as a whole make of data quality, which affects the perspective that data quality is perceived from. Second, task influences the assessment of quality because operational tasks have different needs than strategic tasks. Third, role of the data user affects the quality assessment because different roles in different positions might highlight different aspects in data quality. Fourth, timing affects the assessment of quality because the urgency of a task determines the needs for data. Fifth, individual factors, such as motivation and experience may have an influence on how data quality is assessed. (Even & Shankaranarayanan 2007, 78.)

In their study, Wang and Strong (1996) presented four categories of data quality dimensions to assessing data quality from the user's perspective. The categories included intrinsic, contextual, representational and accessibility data quality. Intrinsic perspective to data quality follows the objective perspective to quality and denotes that the data is essentially correct. The dimensions of intrinsic data quality include accuracy, believability and objectivity. Contextual data quality assesses data quality within the context of the task at hand. The dimensions used in assessing contextual data quality include relevancy, timeliness and completeness. Representational data quality denotes that the data is well presented, interpretable and easy to understand to the data user and thus, includes dimensions of interpretability and representational consistency. Accessibility data quality refers to data that is easily available for all necessary data users but also kept secure from unnecessary data access. Therefore, the dimensions used to assess it include accessibility and access security. (Wang & Strong 1996, 20–22.) The dimension categories by Wang and Strong (1996) are supported by Haug et al. (2009, 1059). However, Haug et al. (2009) argue that accessibility and representational data quality dimensions overlap and hence, they regard representational data quality as being a part of accessibility data quality. Based on this, they created a classification model which includes intrinsic, accessibility and usefulness dimensions. (Haug et al. 2009, 1059.)

Wang and Strong (1996, 22) conclude that data can be considered high quality when it is essentially correct, suitable in the context of the use, expressed clearly, and the data user has access to it. Data user's perspective to data quality is also considered in the definition for data quality by Mosley (2008) who defined data quality as "the degree to which data is accurate, complete, timely, consistent with all requirements and business rules and relevant for a given use". Even and Shankaranarayanan (2007, 83) also note that contextually assessed high quality data can still include flaws but the contextual quality is not negatively affected if the flaws do not interfere with the usability.

In this section, first the definitions for data and quality were presented. This provided the foundation for the definition of data quality. Four perspectives to data quality found in the literature were presented, which included intrinsic, contextual, representational and accessibility data quality. Then data quality dimensions were discussed to broaden understanding about the factors of data quality. This research adopts the user's perspective to assessing data quality. Therefore, in this research the task in which the data is used defines the quality standards that the data must conform to.

2.1.2 Financial data and financial data quality

In this section, the concepts of financial data and financial data quality are introduced. The concepts build on the concepts of data and user's perspective to data quality that were clarified in the previous section.

In this research, the term financial data is used to refer to all data, that directly affects a company's financial processes. Financial in this context does not refer to monetary representation of data but rather to the context in which the data is used. Some researchers refer to this kind of data as accounting data (Warren, Moffitt & Byrnes 2015; Emeka-Nwokeji 2012) or data in accounting system (Bai et al. 2012; Xu 2015). However, because this type of data might also have other purposes in companies besides accounting, the term financial data was adopted for this research to better describe its context. This section combines literature of data used in financial processes.

Financial data is data that is collected and used in financial transactions of a company and it is traditionally used for accounting purposes. Financial reporting is one of the main purposes of financial data which serves a company's internal and external stakeholders. (Warren et al. 2015, 397.) It is mandatory for companies to document their business transactions to financial records as a proof for compliance. Previously these records consisted of physical documentation but today as digital information systems, such as enterprise resource planning (ERP) systems are widely used in accounting and operations, nearly all financial documentation has become digital (Warren et al. 2015, 397; Emeka-Nwokeji 2012, 86). Before information systems, paper documentation did not offer much possibilities for analysis but rather its purpose was to serve as a proof for authorities that business was conducted according to laws and regulations. When information systems were taken into use, it was possible to utilize the data for different purposes, such as planning and evaluating to gain insight of the company's performance (Kaplan, Krishnan, Padman & Peters 1998, 72). Thus, financial data transformed from mere documentation to an important source of business insight and knowledge to companies.

Collecting financial data is mandatory for companies and it is used to serve company's internal as well as external stakeholders. Thus, data quality is a key factor for effective and reliable use of financial data (Bai et al. 2012, 453–454). In the previous section 2.1.1

user’s perspective to data quality was introduced as “fitness for use” which suggests that the requirements for data quality depend on the context where it is used (Kaplan et al. 1998, 57). Also, because the tasks where data is used continuously evolve, contextual understanding of data quality is required (Glowalla & Sunyaev 2014, 668). Therefore, in this research the data user’s perspective is used to assess the requirements for financial data quality. Next the requirements for data quality are discussed according to the tasks of financial data.

The two main purposes in which financial data is used are financial accounting and management accounting. Financial accounting consists of bookkeeping and financial reporting which ensure that a company conducts their business in a compliant manner and it serves primarily external stakeholders. Management accounting uses financial data to evaluate the company’s performance and provide information for decision making. Thus, management accounting serves a company’s internal stakeholders. (Bhimani, Horngren, Datar & Rajan 2015, 3.)

Financial accounting is regulated by law and thus, the requirements for the quality of financial data are mostly defined by the external regulations. The Accounting Act determines the data that companies are obligated to collect and what is required from the data. For example, documents, such as sales invoices must include information of the date, value and the parties of the transaction (KPL 2:5 §). Also, tax regulations determine the tax rates that are used on transactions and thus, transactional data. This means that the data used in financial accounting must accurately represent the business transaction and include all the required information. In other words, the requirements for accuracy and completeness of data that is used for financial accounting are mostly determined by external regulations.

Management accounting produces information for internal use and the quality requirements for financial data are internally defined. Therefore, data quality is assessed as the degree to which the data fulfills the needs of management accounting (Knauer, Nikiforow & Wagener 2020, 100). The information needs of a company determine what kind of data is collected and in which form for it to be utilized and analyzed. Hence, the use cases and analysis needs set the requirements for financial data quality. Financial data used in management accounting includes for example categorizing expenses by different departments of the company. This requires that these internally defined data dimensions, such as department codes, are used accurately and in a consistent manner to produce comparable data for analysis and decision making.

Financial data is commonly stored and used in ERP systems, which also impacts the requirements for data quality. Glowalla and Sunyaev (2014) researched data quality in ERP systems in from the perspective of two tasks: compliance with regulations and using data for analysis. They emphasize that because the tasks have different requirements for data quality, the quality of the data should always be evaluated in the context of the use. They also note that compliance and data analysis tasks overlap to some extent because mandatory reporting also requires analyzing data. (Glowalla & Sunyaev 2014, 670, 677). Because the tasks overlap, this implies that the same underlying data may be used for financial accounting as well as management accounting purposes. As an example, external regulations determine how transactions are categorized to bookkeeping accounts but the account information is also used in management accounting to analyze the cost structure. Therefore, the requirements for the quality of financial data are complex because when the same data is used for different purposes, the data must fulfill the needs of all the use cases. In addition, as use cases change over time and can differ between data users, it poses a challenge to achieving and maintaining high data quality (Wang & Strong 1996, 20; Tayi & Ballou, 1998, 57).

In this section, the concepts of data and data quality in the domain of financial data were clarified. In this research, financial data is defined as the data that directly affects the financial processes of a company. Quality of financial data was defined as the ability of the financial data to fulfill the requirements of the use case. The requirements for financial data quality were then discussed from the perspective of financial accounting and management accounting.

2.1.3 Consequences from poor quality financial data

In this section, the consequences that result poor quality financial data in companies are clarified. First, the defects that consist poor data quality are introduced and then the consequences from poor quality data are discussed in financial accounting and management accounting context.

To understand the consequences of poor data quality, it must first be understood what kind of problems can exist in data quality. There are several kinds of defects that can make data poor quality (Tayi and Ballou 1998, 56). Strong et al. (1997, 104) define a data quality problem as a difficulty on one or more quality dimensions that makes the data unfit for use. Thus, these defects can be discussed according to the data quality dimensions introduced in section 2.1.1 which included accuracy, completeness, consistency and timeliness. Based on the definitions of data quality dimensions by Wang and Strong (1996, 32) a defect in accuracy denotes that the data is incorrect, in other words the value of data does not correspond to the real-life object it represents. Incomplete data occurs when the data is missing certain values that are required by the data user. A defect in data consistency occurs if data is not presented or created in a consistent manner and thus, the data within a data set is not comparable. If the age of the data used for a certain task is not appropriate, a defect in the timeliness dimension occurs.

In the domain of financial data, if high quality data is not available, the goals of financial accounting and management accounting cannot be met (Emeka-Nwokeji 2012, 92). In terms of financial accounting, if the quality of data is poor, the company is not able to meet the financial regulations which risks compliance. The possible consequences for companies from non-compliance are legal sanctions and financial losses. Legal sanctions might be imposed to the company in case of flawed financial statements. (Bai et al. 2012, 453.) If taxes that are reported to the tax authorities include faults, fines might be assigned to the company for misconduct (OVML 37§).

Management accounting provides information for decision making and without high-quality financial data well-informed decisions cannot be made. As Emeka-Nwokeji (2012, 86) states, the financial information is only as good as the quality of that underlying data. Hence, so are the reports and decisions made based on the information. On one hand, if high quality data is not available, the decision makers are unable to base their decisions on facts. On the other hand, if the data is not recognized as poor quality and decisions are made based on that, the decisions might result in outcomes that are not beneficial to the company (Emeka-Nwokeji 2012, 86). Especially reports, such as forecasting and planning which are used in decision making, require high quality data for the company to be able to reliably plan its future operations. It has been observed that high quality data supports business by improved decision making and more efficient business processes (Even & Shankaranarayanan 2007, 75). It can be concluded that the quality of data in accounting systems has a material influence on both external compliance as well as internal decision making (Bai et al. 2012, 453).

In addition to impacting financial accounting and management accounting, poor data quality hinders internal productivity (Emeka-Nwokeji 2012, 87). Cleaning up and correcting the defects in data require extra work that consume the data user's time which leads to increased costs (Haug et al. 2011 173). Furthermore, the time spent on correcting data could otherwise be used in value adding operations. Emeka-Nwokeji (2012, 92) found that improving data quality decreased costs from rework, duplication and unnecessary overhead incurred by poor quality data. Also, wrong decisions made based on erroneous data may lead to decreased efficiency in operations which may result in loss of revenue. In case customer data is incomplete or inaccurate, it has a negative effect on the customer operations which in turn can lead to decreased customer satisfaction. (Bai et al. 2012, 453–454.) In addition, financial data is often used in subsequent systems, such as reporting tools that use the financial data directly from the ERP system. This highlights the importance of the data quality so that the defects in data are not transferred further in the systems and result in false reports. (Glowalla & Sunyaev 2014, 677.)

In this section, the consequences that may result from the use of poor-quality financial data were discussed. It can be concluded that poor quality data interferes effective use of data in financial accounting and management accounting and thus, can lead to financial losses from non-compliance, decision making based on flawed information and ineffective operations. This highlights the importance of high-quality financial data for companies' success.

2.1.4 Barriers to high data quality

As described earlier, poor data quality inhibits effective financial accounting, management accounting and operations in a company. In order to be able to address the defects in data quality, it is relevant to understand what causes the problems. This section will introduce the factors impacting data quality that were identified in the literature. Factors that have a negative effect to data quality are further referred to as data quality barriers. First, data quality barriers identified in general data quality literature are discussed and then the barriers identified in the literature of financial data quality are clarified. Data quality barriers in the general domain were included in this section because literature focusing on the data quality barriers in financial data domain is limited. In this research, the data quality barriers were classified to barriers related to management, roles and responsibilities, communication, and technical barriers. Even though the focus of the research is on organizational factors, technical factors were also included in this section to gain broad understanding of the possible quality barriers.

Silvola, Jääskeläinen, Kropsu-Vehkaperä and Haapasalo (2011) studied challenges for master data quality by conducting qualitative interviews in eight high-tech companies. They found that the most common challenges related to data and processes included unclear master data definitions, poor master data quality, undefined data ownership, incoherent data management practices and the lack of continuous data quality practices. The common challenge identified related to information systems was integrations between the applications. Unclear master data definitions were found to cause problems for communication about data and thus, lead to decreased data quality. Whereas undefined data ownership and incoherent data management practices were found to cause confusion on responsibilities and made data maintenance a laborious task. Regarding the information system, it was observed that when data is transferred through integration, it often negatively affects the quality of data. (Silvola et al. 2011, 155–157.) Silvola et al. (2011, 160) state that their findings imply that the main challenges are not technology related but rather organizational meaning that the issues are not caused by the information systems but rather the practices in data activities among data users.

Haug and Arlbjørn (2010) conducted a literature review covering five articles about data quality barriers. They classified the identified data quality barriers to five themes which included lack of delegation of responsibilities, lack of rewards ensuring valid data, lack of data control routines, lack of employee competencies and lack of user friendliness of the software. Haug and Arlbjørn (2010) validated the themes with a questionnaire of the data quality barriers in master data. They found that barriers that had the biggest impact on data quality were the lack of responsibilities and the lack of control practices. Lack of rewards was found to impact data quality the least. (Haug & Arlbjørn 2010, 296.) Their findings support the research by Silvola et al. (2011) in that unclear responsibilities create a significant barrier for achieving high-quality data. In addition, the findings by Haug & Arlbjørn (2010, 301) emphasize the role of controls. However, they note that defining responsibilities alone may have significant impact on data quality as accountabilities are clear. (Haug & Arlbjørn 2010, 301.) This implies that the data quality barriers are not independent from each other but closely interrelated meaning that a change in one barrier may have an effect on another barrier. The lack of user-friendliness of the software was classified to technical barriers.

O'Brien, Sukumar and Helfert (2013) studied costs of poor data quality and described a case study which had been conducted to identify challenges relating to data quality. The case study was conducted in a large recently privatized company and the main challenge was found to be the low priority of governance practices in the company. Therefore, the employees were unaware of data quality issues and appropriate communication channels and procedure to report data problems were non-existent. Because formal structures and responsibilities regarding data quality were under-developed, the responsibilities around data were not clear to the employees. Also, data activities were conducted on ad-hoc basis which could create a risk to data consistency. Fourthly, the company used local and informal controls which were found to pose a risk to data security and compliance. (O'Brien et al. 2013, 6.) These findings further support the importance of clear responsibilities and defined management practices for data quality.

Tee, Bowen, Doyle and Rohde (2007) examined factors that influence data quality in organizations and collected empirical evidence from a survey and interviews with senior managers at a government-funded service organization. Their findings showed that management responsibilities, such as committing to high data quality, effective communication between stakeholders and awareness of data quality are important factors impacting data quality. (Tee et al. 2007, 351.) It is noteworthy that in their study only one of nine hypothesis regarded information systems and thus, the focus of the study was on organizational factors.

Data quality barriers can also be identified in the domain of financial data. Fletcher, Robbert, Mohamad and Middleton (2005) studied assessment and improvement of financial data quality by using a process approach. They found that communication and knowledge sharing between the people working with data are fundamental for achieving high-quality data. This is because knowledge of the collection, storage and utilization process of data enhances the data creators understanding of the requirements for data and thus, contributes to creation of high-quality data. (Fletcher et al. 2005.) Therefore, the lack of communication can be translated into a possible data quality barrier.

Glowalla and Sunyaev (2014) conducted a qualitative research on the interdependency of an ERP system and data quality. They conducted interviews in insurance companies, which used ERP systems mainly in accounting. Based on their findings, Glowalla and Sunyaev (2014) share the view that the lack of data quality management is a key challenge to data quality in organizations. This is because without intentionally managing data quality, companies might adopt external regulations as their requirements for quality and ignore their organization-specific needs for data. In addition, management is needed to ensure that the context-sensitivity of data is taken into account if the data is later used in a different context than it was created for. (Glowalla & Sunyaev 2014, 679–680.)

Knauer et al. (2020, 102–103) investigated the determinants of information system quality and data quality in management accounting and found that internal IT knowledge is positively correlated with data quality. This means that if the company does not have adequate IT competencies, it can have a negative impact on financial data quality. Unskilled staff are not as capable to adapt the information system to organizational needs as effective use of data would require. Knauer et al. (2020, 102) add that as data today is stored and used in information systems, internal IT competencies need to be developed in all organization’s functions that operate with data, not only in the IT function. (Knauer et al. 2020, 102–103.)

Emeka-Nwokeji (2012) conducted a research to improve data quality in accounting information systems. She states that most of data quality problems in accounting data are due to the lack of effective control system for evaluating the correctness of data. Her survey results indicated that data quality management practices lead to decreased costs and improved organizational performance. (Emeka-Nwokeji 2012, 86, 92.) This supports the argument that lack of data quality management practices create a data quality barrier.

Xu (2015) studied the most important factors influencing data quality in accounting information systems. She conducted an extensive literature review of critical success factors for data quality, which included i.a. management commitment, training and education, organizational structure and culture, change management and controls. Then, a survey was conducted in accounting information system context to identify the most important success factors for data quality in accounting information systems. According to the findings of her empirical study, the most important factors affecting data quality were management commitment, nature of accounting information systems and input controls. Management commitment refers to that the importance of data quality is recognized by management. Nature of accounting information system denotes the suitability and ease of use of the system. Input controls aim to prevent input errors in data entry. (Xu 2015, 9, 13.)

The data quality barriers identified from the literature were classified to four themes which included management, roles and responsibilities, communication and technology related data quality barriers. A barrier was classified as management related if it regarded practices with data that do not support high data quality and thus, lead to data quality problems. The category of roles and responsibilities includes barriers which concerned the division of decision-making authority and responsibility areas regarding data. Unclear roles and responsibilities may cause confusion over decision making and conducting tasks and thus, negatively affect data quality. Barriers that were classified as communication related are factors that hinder effective communication among data stakeholders. The lack of employee competencies was also classified as communication related barrier because training and education can be regarded as a form of communication and thus, the lack of competencies is the result of inadequate communication. As Fletcher et al. (2005) found, communication and knowledge sharing are a perquisite for achieving high data quality and therefore, weak communication leads to poor data quality. Barriers that concerned information systems were classified as technology related barriers. The data quality barriers identified in the literature have been summarized in the table 1 below according to the mentioned themes.

Table 1. Data quality barriers

Theme	*Data quality barrier*	*References*
Management	Incoherent data management practices	Silvola et al. (2011)
	The lack of continuous data quality practices	Silvola et al. (2011)
	Lack of rewards	Haug & Arlbjørn (2010)
	Lack of control practices	Haug & Arlbjørn (2010); Emeka-Nwokeji (2012); Xu (2015)
	Low priority of governance practices	O'Brien et al. (2013)
	Local and informal controls	O'Brien et al. (2013)
	Lack of data quality management	Glowalla & Sunyaev (2014); Emeka-Nwokeji (2012)
	Low management commitment	Tee et al. (2007); Xu (2015)
Roles & responsibilities	Undefined data ownership	Silvola et al. (2011)
	Lack of responsibilities	Haug & Arlbjørn (2010); O'Brien et al. (2013)
Communication	Lack of communication	Tee et al. (2007); Fletcher et al. (2005)
	Unclear data definitions	Silvola et al. (2011)
	Lack of communication channels for data quality issues	O'Brien et al. (2013)
	Lack of awareness of data quality	Tee et al. (2007)
	Lack of employee competencies	Haug & Arlbjørn (2010); Knauer et al. (2020)
Technology	Integrations between applications	Silvola et al. (2011)
	Lack of user-friendliness of the software	Haug & Arlbjørn (2010)
	Nature of the accounting information system	Xu (2015)

In the literature several methods are suggested to address the quality problems in financial data which are grouped to data-focused and process-focused strategies. Data-focused strategies aim to improve the existing data, such as replace an incorrect value in data with a correct value, whereas process-focused strategies aim to develop data processes to affect the root causes that lead to data quality problems. (Liu, Feng, Zhao & Wang 2020, 2.) Even though the identified data quality barriers imply that most quality challenges are organizational indicating that process-focused strategies would yield the best results, it must be noted that data-focused strategies are also used in several studies conducted on improving financial data quality. For example, Du and Zhou (2012) presented an ontology-based framework for improving the quality of financial data especially focusing on data inconsistency problems in financial data. Alpar & Winkelsträter (2014) took a data-focused approach and developed a method to identify possible defects in the quality of financial data. This implies that technical solutions have also been found beneficial for improving financial data quality.

According to Cao and Zhu (2013, 3), research shows that improving data processes and influencing the root causes of data quality problems are an effective way to improve data quality. Examples of suggested process-focused methods to improve financial data quality include a research by Bai et al. (2012) who proposed a decision model to achieve optimal level of control in data entry to manage data quality risks in accounting information systems. Also, Liu et al. (2020) followed the process-focused strategy and developed a model to allocate resources to data quality improvements in a cost-effective manner focusing on controlling behavior in data operations.

In this section, the factors influencing data quality called data quality barriers were discussed to understand the causes of data quality problems. The identified data quality barriers imply that most of the factors influencing data quality are organizational rather than technical. Therefore, in order to influence data quality, the focus should be placed on the organizational processes regarding data. However, also data-focused methods are suggested in the literature for improving data quality which implies that data quality can be also addressed with technical solutions. In this research the focus is on addressing the organizational data quality barriers.

2.2 Data Governance

In this chapter, the concept of data governance is clarified. First, the definition for data governance is introduced. Then the different principles of data governance are clarified and finally, the antecedents and consequences of data governance are discussed.

2.2.1 Data governance and IT governance

In this section, first the definition for data governance is clarified. Then the development of data governance and the role of IT governance as the foundation of data governance are discussed.

The existing literature does not offer a standard definition for data governance and the existing definitions have slightly different emphasis. Mosley (2008) defines data governance as “the exercise of authority, control and shared decision making (planning, monitoring and enforcement) over the management of data assets” (Mosley 2008.) According to Weber et al. (2009, 6) “data governance specifies the framework for decision rights and accountabilities to encourage desirable behavior in the use of data”. Khatri and Brown (2010, 150) define data governance as the assignment of decision rights and accountabilities for an organization’s decision making about its data assets. Cheong and Chang (2007, 1001) state that data governance “defines policies and procedures to ensure proactive and effective data management”. Otto (2011a, 241) defines data governance as the “allocation of decision-making rights and related duties in the management of data”.

The mentioned definitions outline the key elements of data governance, even though some variation can be found in them. Firstly, several of the definitions mention the goal of managing data or data assets. Secondly, the definitions refer to assigning decision-making rights over data. Thirdly, some definitions also include the allocation of responsibilities, accountabilities or duties regarding data. Cheong and Chang (2007, 1001) also add policies and procedures as a part of data governance which guide data activities in companies. These summarize the key elements of data governance. Because data activities are conducted by employees, the data assets can be managed by governing the actions of the employees.

Data governance as a concept has been developed during the past ten years (Alhassan et al. 2019, 989). The theory of data governance is built on IT governance which refers to the framework of decision rights and accountabilities for decision making regarding organization’s IT assets (Khatri & Brown 2010, 149). The difference between data governance and IT governance can be understood by comparing the definitions of data governance and IT governance by Khatri and Brown (2010, 149–150). Where IT governance is concerned with the decisions on IT assets, data governance in concerned with the decisions on data assets. From the perspective of data governance, IT is seen as the infrastructure where data activities take place (Cheong & Chang 2007, 1002).

The literature of IT governance has an emphasis on the physical IT artifacts, such as software and hardware and their management and control (Tallon et al. 2013, 143). That is understandable due to the history of the physical IT infrastructure in the key focus of information governance. However, as the amount of data has grown exponentially in the past years, it has raised the need to include also nonphysical aspects, such as data, to this discussion (Tallon et al. 2013, 142, 148). Kooper, Maes and Lindgreen (2011, 196) note that the limitation of IT governance is that IT is not concerned with the creation of information but that its focus is on managing the resources. As stated earlier, data is not included in IT but rather IT provides the necessary resources for the data activities. Because data and IT are independent from each other, data cannot be controlled with IT governance tools. As data is created and used by people, it is affected by human factors, which needs to be considered when aiming to govern it. Weber et al. (2009, 1–2) state that organizational issues have been ignored in data management and data quality domains. Thus, they combined IT governance theory and organizational theory in their research on data governance to fill this gap in literature. (Weber et al. 2009.)

From a practical point of view data governance has two goals. First, to ensure that the quality of data is on a good level for decision making. Second, to make sure that misuse or human error do not cause harm to the data quality and thus to the value of data. (Tallon et al. 2013, 142.) Otto (2011b, 7) describes data governance goals in more detail. According to his research the goals pursued with data governance include ensuring compliance, supporting decision-making, improving customer satisfaction, increasing operational efficiency, supporting business integration and increasing data quality (Otto 2011b, 7). Otto (2011b, 7) lists these goals as formal goals which are the result of effective data governance and thus, can be achieved when the data is high-quality. When the data is high quality, the company does not need to worry about compliance regarding data. High-quality data enhances decision-making which may lead to increased customer satisfaction and increasing operational efficiency due to right actions in these areas. Thus, by ensuring high data quality, data governance contributes to success of data activities.

The goal of achieving high data quality is also emphasized in the findings of Pierce, Dismute and Yonke (2008, 5) who researched the state of data governance in companies. She found that the leading goals that companies aimed to achieve with data governance were improving data quality, facilitating business intelligence activities and addressing compliance issues (Pierce et al. 2008, 5). According to Kooper et al. (2011, 195) companies that have implemented data governance are able to collect, analyze and use data more effectively than companies without data governance. Hence, those companies can derive more value from their data assets, which benefits the company and supports its objectives. (Kooper et al. 2011, 195.) Therefore, data governance builds a foundation for effective use of a company's data assets (Todd 2008, 30).

In this section, the concept of data governance was outlined. The key elements of data governance were identified from the definitions, which include the purpose to manage data assets, division of decision-making rights and responsibilities as well as definition of practices to ensure effective data management. Lastly, the goals of data governance were described of which ensuring high data quality was seen as the key goal.

2.2.2 Data governance roles and principles

This section clarifies the data governance structure. As described above, data governance is a framework for governing data activities to ensure desirable behavior with data. Because data is affected by human behavior, data governance aims to influence the data at its roots. To achieve this, there are two main data governance activities. First, data governance defines roles for people working with data and assigns accountabilities. Second, data governance establishes company-wide guidelines and standards regarding data that are aligned with company's strategy. (Weber et al. 2009, 6.) Roles and accountabilities define the structure of who can make decisions and who are responsible for which areas. Guidelines and standards aim to facilitate common understanding of data in a company and encourage behavior that has contributes to high data quality.

The three common roles mentioned in the literature include data owner, data steward and data committee (Otto 2011a, 242). Typically, data owners are responsible for the data assets in their business function (Abraham et al. 2019, 428) and provide the standards for data quality (Otto 2011ba, 242). Data owner should be most aware of the business requirements of certain data and ensure that these requirements are met. Data steward works together with the data owner. Whereas data owner provides the guidelines, the responsibility of the data steward is to execute data activities according to these guidelines. In other words, data steward is responsible for the actual data management. (Otto 2011b, 8.) Data stewards are often knowledgeable of the business requirements as well as IT aspects in order to communicate with data users and translate technical needs to IT (Cheong & Chang 2007, 1005). Typically, data stewards are responsible for data from a certain department and support the data users in the department to ensure desired behavior in data use. In addition, data stewards evaluate the requirements of data and solve problems encountered with data. (Otto 2011a, 242.)

Data committee, which is the central decision-maker in an organization, establishes organization-wide governance practices for data (Otto 2011a, 242). A data governance council is another name used for a data committee and its responsibility is to balance the interests of the different stakeholders when company-wide decisions are made (Cheong & Chang 2007, 1004–1005). Data governance council prevents making siloed decisions when appropriate stakeholders are included in decision making. According to Otto (2011b, 8) data governance council consist of data owners whereas Wende (2007, 421) describes that data stewards attend in data governance board meetings. Data sponsor is also identified in the literature (Otto 2011b, 8; Weber et al. 2009, 11). Data sponsor is described as an executive level role whose responsibility is to give a mandate and provide strategic direction for data governance activities in an organization (Weber et al. 2009, 11).

Some researchers also add data producer and data consumer to the categories (Abraham et al. 2019, 429; Cheong & Chang 2007, 1005). Data producer refers to the person who creates data and data consumer refers to the person who utilizes the data for example for reporting or analysis (Abraham et al. 2019, 429). Because data producer and data consumer are often those, who work with creating and utilizing the data the most, it is also important to define their roles and responsibilities regarding data. Data users report data related problems and needs that they encounter in data use (Cheong & Chang 2007, 1005). It is important to note that the roles described in the literature are examples of how roles and responsibilities can be divided in companies, but a company should always design its own data governance roles according to its individual needs. (Weber et al. 2009, 9).

Consistency is one of the most important factors for achieving high data quality. Thus, the purpose of data governance principles is to facilitate consistent behavior in all data activities. To achieve this, everyone working with data should have common understanding of data and use the same definitions for data company wide. By implementing company-wide practices, policies and norms for actions around data that are aligned, data governance aims to contribute achieving common understanding of data in the organization. (Weber et al. 2009, 6.) In this vein, in the research by Pierce et al. (2008, 19) standardizing data definitions across the organization was found to be the most mentioned governance activity in the interviewed companies.

Alhassan et al. (2019) studied critical success factors of data governance in their research by interviewing the employees of a case organization. They identified seven critical success factors which outline the core of successful data governance framework. The factors included employee capabilities, clear processes, flexible tools and technologies, standardized easy-to-follow policies, clear roles and responsibilities, clear requirements and a tangible strategy. (Alhassan et al. 2019, 104, 108.) This implies that data governance should be developed to support these factors in an organization. Alhassan et al. (2019) also investigated the relationships between the success factors and found that they are often interrelated. For example, even if data governance policies were well designed and carefully implemented, problems with data will persist unless also roles and responsibility areas are clear in the organization. Thus, successful data governance requires that all of the critical success factors are taken into consideration. (Alhassan et al. 2019, 106, 108.)

In this section, the key data governance activities were introduced which are defining roles and responsibilities and establishing company-wide practices for working with data. The roles mentioned in the literature were discussed to gain understanding of the division of responsibilities. Lastly, the success factors of data governance were introduced. The success factors imply that data governance activities are not independent from each other but rather they are interrelated and one factor may have an effect on another factor.

2.2.3 Antecedents and consequences of data governance

In this section, the antecedents and consequences of data governance are introduced. Antecedents of data governance are factors that predict the implementation of data governance practices and may either facilitate or limit their adoption in an organization (Tallon et al. 2013, 143). They can be classified to external and internal antecedents. External antecedents concern the requirements in a company's legal and regulatory environment that impact the data usage and storing. Internal antecedents are a company's internal factors which include company culture, strategy and organizational structure. (Abraham et al. 2019, 432.) For example, industry regulations as an external antecedent may increase the need for a company to ensure compliance, which may drive the adoption of data governance. Company culture is perceived as an internal antecedent and especially data-driven culture may drive the adoption of data governance to ensure secure and effective use of data. In addition, internal IT knowledge was found to facilitate the implementation of data governance whereas product complexity and unfit systems were found to inhibit data governance. (Tallon et al. 2013, 159–161).

The consequences of data governance are an important research area because the main argument for implementing data governance are the benefits that result from it. Also, the researchers agree that by implementing a data governance framework, companies can improve the management and thus, the quality of their data (Cheong & Chang 2007, 999). Tallon et al. (2013, 166) studied the consequences of implementing data governance in organizations and found several benefits that organizations had grasped. Even though the results were affected by the industries where the participated companies operated in, common improvements that were found included enhanced decision making, reduced costs and higher customer satisfaction. (Tallon et al. 2013, 166.) Abueed and Aga (2019, 10) found that implementing data governance resulted in improved decision making and knowledge creation in a company. They argue that managers can enhance knowledge creation in companies by fostering data governance policies. (Abueed & Aga 2019, 11.) Kamioka, Luo and Tapanainen (2016) investigated the effect of data governance on marketing performance. They found that the accountabilities in data governance enhance data utilization which consequently has a positive effect on marketing performance. (Kamioka et al. 2016, 7.)

An important observation by Tallon et al. (2013) was that the results of data governance do not continuously improve when more data governance activities are adopted. Rather, over-governance can cause disadvantages, such as hindered data-driven innovation. Hence, the appropriate level of data governance must be assessed according to the needs and goals of the organization in question. (Tallon et al. 2013, 167.) The optimal level is achieved when the external requirements are followed but the employees have adequate room to conduct their data activities (Kooper et al. 2011, 197).

In this section, the data governance antecedents and consequences were clarified. Antecedents impact the adoption of data governance and may facilitate or inhibit the adoption. Therefore, it is important for companies to identify their individual antecedents in order to consider them in the implementation process. The consequences of data governance imply that data governance is an effective tool for enhancing data quality and data use in organizations when an optimal level of governance is achieved.

2.3 Data Governance Models

In this section, two data governance models will be introduced. First is the data governance model by Khatri and Brown (2010) and then the contingency model by Weber et al. (2009).

Data governance model by Khatri and Brown

Drawing on IT governance theory, Khatri and Brown (2010, 149–150) developed a data governance framework consisting of five decision domains: Data principles, Data quality, Metadata, Data access and Data lifecycle. These decision domains are next introduced.

Data principles form the basis for data governance in an organization. The principles cover the top line rules and norms, according to which the data governance activities are organized. Therefore, data principles should be aligned with the business needs for data as they create the guidelines for the encouraged behavior with data in the organization. Also, legislative and regulatory compliance need to be taken into consideration when creating data principles as these principles found the basis for the four other decision domains. The basis for setting data principles for an organization is the understanding of data as an asset that needs to be managed in order to grasp its potential value (Khatri & Brown 2010, 150).

Data quality as a concept was discussed more in detail in section 2.1.1. This is because data quality is considered a fundamental element of a data governance program (Alhassan et al. 2019, 102) and thus, the concept was separated to discuss it in more detail. Following the user's perspective to quality, Khatri and Brown (2010, 150) define data quality as the extent to which the data meets the user's requirements. They argue that data quality can only be assessed on the different quality dimensions, such as timeliness, in regard to the use case at hand. The link between data governance and data quality is that the data quality standards and requirements are defined in data governance. Also, data governance assigns the responsibility and decision-making rights regarding data quality decisions which are essential to ensure effective data management. (Khatri & Brown 2010, 150.)

Metadata is commonly referred to as "data about data", in other words information about certain data. Metadata includes descriptions of the characteristics of the data and defines standardized meanings for different data objects. The aim is to provide a common company-wide language of data to ensure that everyone interprets the data in the same way. (Khatri & Brown, 150.)

Data access decisions refer to the policy of rights to for employees to access a certain organization's data. The main driver for establishing data access policies is the regulatory environment of an organization. The objective of data access policy is to is to protect the confidentiality and integrity of the data as well as provide availability to the relevant data for the right people. (Khatri & Brown 2010, 151.)

Data does not usually stay unchanged in an organizations' IT architecture, but rather it goes through different stages which form the data lifecycle. These stages include creating, storing, processing and analyzing. The data lifecycle is important to manage in order to have a holistic understanding of the organization's data. This includes mapping and defining the existing data of an organization because in order for a company to manage its data, it needs to know what kind of data it stores. A well-managed data lifecycle can result in improved data storage use and thus, decreased costs. (Khatri & Brown 2010, 151.) The decision domains are summarized in the figure 2 below.

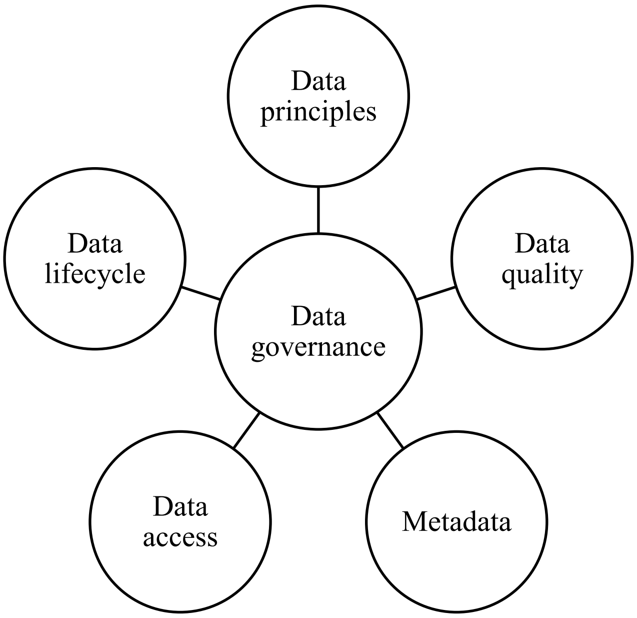

Figure 2. Decision domains of data governance (based on Khatri & Brown 2010)

Contingency Model by Weber et al. (2009)

Even though data governance has been found effective in improving data operations and data quality, companies tend to face challenges in implementing it. Weber et al. (2009, 7–8) suggest that a possible reason for the challenges that companies face with data governance is that the literature offered only a one universal way of applying it to companies which does not take into account the different contingency factors that affect its implementation. Also, Wende (2007, 418) notes that in the academic as well as practitioners' literature data governance is often misunderstood as a universal approach implying that the same framework can be implemented to all companies (Wende 2007, 418).

Therefore, Weber et al. (2009, 3) suggest the opposite arguing that there is no common approach for all companies to data governance but rather that the company specific factors determine the individual data governance framework for each company to best serve their needs. Weber et al. (2009) developed a contingency model to data governance by identifying seven contingency factors that affect the composition of data governance. Contingency factors mean certain characteristics that are individual for each company and differ among companies. The contingency factors they identified are performance strategy, diversification breadth, organization structure, competitive strategy, degree of process harmonization, degree of market regulation and decision-making style. These contingency factors need to be taken into account when designing data governance to ensure effective data quality management and thus, high data quality. (Weber et al. 2009, 17.)

The contingency factors affect especially the structure of the roles and responsibilities in the data governance framework. This is because the allocation of roles and responsibilities is influenced by whether the contingency factors require centralized or decentralized actions. For example, if the organization structure is centralized in the company, data governance framework is created in a centralized style to support the organization structure. In their study Weber et al. (2009) found that the contingency factors mediate the fit between the data governance framework design and successful data quality management (Weber et al. 2009, 19–23). Figure 3 of the contingency model below presents the relationship of the data governance design and data quality management success, which is affected by the seven different contingency factors.

Figure 3. Contingency model to data governance (modified from Weber et al. 2009, 17)

2.4 Summary

The second chapter covered the theoretical background of the research by introducing the key concepts of the research, which were data quality and data governance. This section is a summary of the theoretical framework that synthesizes the main concepts of data quality and data governance and aims to provide answers to the research questions based on the theoretical framework.

To understand data quality in the context of financial data, this study adopted the user's perspective to data quality which means that data quality is assessed as the ability of the data to fulfill the requirements of the data user and the task it is used for (Wang & Strong, 1996, 6). In the context of financial data this means that the different tasks of financial data set the requirements for data quality. The requirements for data used in financial accounting are mostly determined by external regulations. Whereas the requirements for data used in management accounting are determined internally according to a company's information needs. When the same data is used for both financial and management accounting purposes, the requirements for data quality are even higher for the data to be able to fulfill the requirements of both of the tasks. If the quality of financial data is not on an adequate level for usage, it can have negative effects for the company. Consequences from poor data quality include legal sanctions from non-compliance, decision making based on misleading information and inefficient operations. (Bai et al. 2012, 453–454.) Therefore, high quality financial data is essential for a company's success.

The first research question focuses on examining the possible challenges that exist in financial data quality. There are several reasons that might lead to poor quality data, which were introduced as data quality barriers in this research. The data quality barriers that were identified from the general data quality literature and financial data quality literature were categorized to four themes which included management, roles and responsibilities, communication and technology related data quality barriers.

Silvola et al. (2011, 160) state that problems with data quality are mostly organizational and process based instead of technical. Because management, roles and responsibilities and communication related barriers deal with human-related factors, they can be regarded as organizational barriers. Whereas technology related barriers are concerned with the information system characteristics. Data users create, store and process data, and hence, their actions impact data quality (Kooper et al. 2011, 197). As the barriers indicate, if the responsibilities and practices among data users are not clear, it will negatively influence data quality (Haug & Arlbjørn 2010, 301). Therefore, in order to address the organizational barriers, the practices and behavior among the data users need to be influenced. Technical data quality barriers cannot be influenced with organizational practices but rather they require development of the software in which data is used.

The findings of Watson, Fuller & Ariyachandra (2004) also support the observation of the great impact that organizational factors have on data quality. In their study Watson et al. (2004) investigated the adoption of technical data management systems as a method to improve the quality of an organization's data. They found that organizational factors, such as management participation and education have a key role for the system implementation and thus, achieving high data quality. This implies that the issues with data management cannot be addressed with only technical solutions. (Watson et al. 2004, 436, 447.) Glowalla and Sunyaev (2014, 679) also emphasize the importance of organizational factors in data quality. They concluded in their study that achieving high data quality requires implementing data quality management structures and the key step in that is determining owners for the data (Glowalla & Sunyaev 2014, 678). Therefore, the organizational issues need to be addressed by solutions affecting the organizational factors.

Cheong and Chang (2007, 1001) state that in order to address the quality issues in data, data needs to be governed. Also, according to Benfeldt Nielsen (2017, 120) data governance is seen as a promising way in addressing organizational data challenges due to which it has been a growing interest of academics as well as practitioners. This is because achieving high data quality requires that data is well managed and clear practices for working with data exist among data stakeholders. Hence, the concept of data governance was developed to address the issue of poor-quality data. Instead of correcting defects of data in the information system, data governance aims to affect the root cause of the data quality problems and hence, mitigate the data quality barriers. Data governance works as the framework that determines how data management and the daily operations around data are organized. By organizing data activities, data governance aims to ensure that data assets are properly managed. This is done through two main tasks, which include allocating decision-making authority and responsibility areas and establishing practices that support desired behavior in data activities. (Weber et al. 2009, 6.)

The second research question was set to examine how data governance can be utilized to address data quality challenges. Because data governance addresses organizational barriers, the barriers related to management, roles and responsibilities and communication are next reflected on data governance.

Management related barriers included inadequate management practices for data, lack of rewards, weak controls, and low management commitment which inhibit achieving high data quality. Unclear practices often lead to inconsistent ways of working, which may cause defects to data consistency and accuracy. To address these barriers, data governance framework includes organization-wide practices to ensure common ways of working with data. Firstly, data governance sets data quality as a priority and emphasizes the need to manage a company's data assets. Some researchers suggest establishing data strategy to highlight the strategic importance of data assets (Weber et al. 2009, 10; Alhassan et al. 2019, 107). Secondly, data management processes should be defined because they determine how data objects are maintained, used and deleted. (Weber et al. 2009, 11–12). Clearly defined processes ensure consistent treatment of data objects over their lifecycle to maintain data quality. Clear data processes and procedures were also recognized as a critical success factor of data governance (Alhassan et al. 2019, 104–105). In addition, control practices may be established to monitor the performance of management practices (Weber et al. 2009, 10). Data policies further contribute to common practices in data activities by briefly stating high-level guidelines and norms for working with data. Principles and policies included in data governance framework ensure that everyone conducts data activities according to the same guidelines and thus, supports achieving high quality data. Establishing management practices require that roles and responsibilities regarding data are divided. (Alhassan et al. 2019, 106–107.)

Roles and responsibilities related barriers included undefined data ownership and lack of delegation of responsibilities. Undefined roles may cause confusion among data users and create a risk for data quality as nobody is accountable for ensuring high data quality. Data governance addresses these barriers by specifying a structure of decision-making rights and responsibilities to contribute to desired behavior in data activities (Weber et al. 2009, 6). This means that people working with data are recognized and suitable roles are assigned to them. In order to clarify the responsibilities related to data, the responsibilities of each role are defined and documented (Alhassan et al. 2019, 106). Decision-making authority is defined among the different roles to clarify who is authorized to make decisions regarding data. This ensures that decisions are made to benefit the data assets and data quality. Depending on the needs of the company, they might adopt a centralized or a decentralized approach for placing decision rights. In a centralized approach, the decision-making right is placed on one subject, such as the data governance board whereas in a decentralized approach decision rights are placed on data stewards. (Weber et al. 2009, 14.) If different interests exist in the organization, a data committee may be established to balance conflicting views (Otto 2011a, 242).

The data quality barriers related to communication included lack of communication, unclear data definitions, awareness of data quality and lack of competencies. Because communication has a significant impact on data quality, facilitating effective communication has also been viewed as an important goal in data governance. The key step towards achieving effective communication is facilitating common understanding about data in the organization. Thus, an important data governance task is to develop unambiguous definitions for data objects (Weber et al. 2009, 12) which Khatri and Brown (2010, 150) refer to as metadata. This ensures that the data stakeholders have common understanding of different terms, which contributes to effective communication. Without clear and standardized definitions, employees working with data might interpret it differently which may result in misunderstandings and decreased data quality. Because employee competencies affect the conducting of data processes, they have an impact on data quality. Therefore, training employees about data governance issues and data policies is important for increasing the awareness of data quality among data stakeholders (Alhassan et al. 2019, 104). Data governance aims to ensure that data stakeholders have adequate knowledge of data issues and common understanding of data definitions and thus, can effectively communicate about data. Table 2 presents a summary of the data governance activities for addressing organizational data quality barriers

Table 2. Summary of data governance activities for addressing data quality barriers

Management	Lack of management practices	Establishing clear practices and policies for working with data
	Lack of controls	Introducing necessary level of controls
	Lack of rewards	Developing practices to encourage desired behavior with data
	Lack of management commitment	Developing a data strategy to view data as a strategic asset
Roles and responsibilities	Undefined data ownership and responsibilities	Defining roles, decision-making authority and responsibilities regarding data
Communication	Lack of communication	Facilitating common understanding to increase communication
	Unclear data definitions	Developing standardized definitions for data, establishing metadata
	Lack of awareness	Clear communication of data quality requirements
	Lack of competencies	Education and training

According to the contingency model of data governance presented by Weber et al. (2009) data governance should always be developed individually for each organization's specific needs in order for data governance to provide the expected results. This contingency model can be adapted to the context of financial data. This means that the requirements for data quality are defined by the context of financial data and thus, the requirements for financial data must be considered in the data governance framework. This is done by evaluating the tasks and defining the requirements for data quality for the tasks of financial data. The data quality requirements further shape the structure of the data governance framework along with the contingency factors.

3 EMPIRICAL RESEARCH

The third chapter introduces the choices that were made to conduct the empirical research. First, the case company as the research subject is introduced to provide a context in which the empirical data was collected. Second, the research approach following the design-based research is clarified. Then, data collection and analysis methods are presented and finally, reliability and validity of the research are discussed.

3.1 The case company

The case company of this research is Framery Oy. Framery Oy is a Finnish company that is the pioneering manufacturer of acoustic phone booths and private spaces for open offices. The company operates in the office furniture industry and the most of its business takes place in the business-to-business market. (Framery 2021.) The business idea was developed by the two company founders who wanted to create a solution for noise and privacy issues in open offices. The purpose of the company is to make office employees happier and more efficient by minimizing sound distractions in office spaces. Framery's product offering includes 4 products, which are acoustics spaces of different sizes and features. (Framery 2020.)

Framery was founded in 2010. After a slow start, the company was able to raise awareness within the furniture industry, which led to a fast growth in sales. Framery nearly doubled its revenue each year from 2015 to 2019, during which the revenue grew from 5 million to 103 million. In 2019, Framery was ranked as 20th in the Financial Times FT1000 List of Fastest Growing Companies. (Framery 2020.) The accelerating growth provided Framery a great opportunity to capture market share but simultaneously put pressure on the daily operations. At the time, the priority in the company was to be able to meet the market demand which led to compromises in other areas. During the heavy growth, the management style was typical for a start-up with lean and light operative management leading to ad-hoc decision-making style. Also, processes were not developed because they were constantly changing. The company successfully managed to keep up with the growing sales but the lack of structure is still visible in some of the company's operations. By today, the company has grown from a startup to a middle-sized company and thus, structure is needed to ensure stability and consistency in the operations.

Framery's finance department consists of financial accounting and management accounting functions which both consist of own teams. Despite the function separation, the teams work closely together. The fast growth has also resulted in frequent changes in the finance department during the past few years. The financial accounting team currently consists of 11 employees, of which 5 were recruited during the year 2020. Management accounting team also grew from 3 to 6 employees during 2020 when new roles in the team were established. Due to this, the teams are still quite young in their current structures and the ways of working are still being formed.

3.2 Research approach

This research follows the design-based approach which was briefly introduced in the section 1.3 Research methodology. The goal of a design-based research is to create a solution for a single organization's problem. A design-based research is initiated when a company identifies a need or a problem that they want to resolve. If there is not a ready solution that can be implemented, a design-based research can be started to develop a solution specifically for the organization's needs. Creating a solution for the organization requires tight cooperation between the researcher and the actors in the organization. The researcher combines theories from the literature with the situational theory in order to form an organization-specific theory that is the basis for the developed solution. (Tamminen 1993, 154–157.) Tamminen (1993, 158–161) identifies four stages that form a design-based research process which include familiarizing oneself with the organization, creative thinking, evaluation and commitment.

To achieve understanding of the organization's problem, the researcher must get familiar with the case organization and its problem (Tamminen 1993, 159). In this research, preliminary discussions about the company's needs were held with the finance director in September 2020. The discussions gave the researcher the initial understanding of the needs that the company was seeking a solution for. The case organization had recognized the need to ensure the quality of financial data to be able to support the business needs. It was found important that processes with financial data were developed to provide high quality financial data in the future and to minimize data quality risks. Thus, this was set as the initial goal of the research.

The discussions with the finance director provided the researcher with the necessary information to start building the theoretical framework. The theoretical framework was built on data quality literature, especially focusing on the causes for data quality problems to increase the researcher's understanding of the topic. Based on the theory of data quality barriers, data governance was chosen as the primary solution for addressing these problems.

When the researcher had acquired adequate knowledge of the data quality problems in the literature, it was relevant to gather more information about the case organization's problems. Interviews were conducted with the most important stakeholders of financial data in the company to gather more detailed information about the challenges that they encounter with the quality of financial data in their work. The interviewees were also asked to describe the current processes with financial data to increase understanding of their current state in the company. In addition, the interviewees reflected their own needs and development ideas regarding the current processes. The information gathered in the interviews was used to answer the first research question: What kind of challenges regarding the quality of financial data exist?

The interview data was used to classify the data quality challenges in financial data. These findings were then used to validate data governance as the solution for the company's needs. The information gathered from the interviews was combined with theories found in the literature to develop the company-specific solution for the case-organization's problem with financial data quality. Therefore, a data governance framework was developed for the case company. The development of the company-specific solution aims to provide answer to the second research question: How can the data quality problems in financial data be addressed with data governance?

3.3 Data collection and analysis

This section clarifies the choices that were made in the data collection and analysis in order to collect relevant data for the research purpose. Data collection methods are often determined by what kind of information the researcher is interested in finding (Hirsjärvi et al. 2009, 179). As was clarified in the section 3.2 this research follows the design-based research approach. In order to create a solution for the company's problem, the researcher must first gather information of the company and the company's problem. Sources for information may include company documents, interviews, and observation. (Tamminen 1993, 159.)

In this research, the empirical data was collected from interviews in the case company. Due to the qualitative nature of the study, interviews were found to be a suitable data collection method because they enabled gathering detailed information and gaining understanding about the data quality problems. Because the hermeneutic approach of this study emphasizes people's interpretations as a source of information, interviews enabled collecting information from the employees' perspective. Interviews as a data collection method also support the user's perspective to data quality used in this study. It denotes that data quality is assessed from the data user's viewpoint and thus, the data user is the best source of information for understanding data quality.

Interview types are traditionally classified to three groups according to the level of structure in the interview situation, which include structured, semi-structured and unstructured interview. On one end, structured interview is often conducted as a survey where the questions and the research structure are standardized. Structured interview fits well to a research where the researched sample is large and thus, the findings can be generalized to other contexts. On the other end, unstructured interview is conducted in an open style without ready-made questions and the interview is led by the interviewees' thoughts and opinions. Unstructured interview resembles a discussion and it is used to collect highly in-depth information from very few interviewees. Semi-structured interview is the combination of structured and unstructured interview types. The questions and the arrangement of the interview are not standardized but the interview follows the themes decided by the interviewer. Semi-structured interview provides in-depth information about the researched theme but it is flexible and can be adjusted according to the interview situation. (Hirsjärvi et al. 2009, 203–204.)

The interviews in this research followed the semi-structured approach. Semi-structured interview was chosen because gaining broad understanding of the data quality challenges in the case company required several data users to be interviewed. In addition, open form answers enabled the interviewees to present their own experiences without ready-made alternatives. The flexible form of interview also enabled the interviewer to ask additional questions when more detailed information was necessary. Structured interview would not have provided the needed in-depth information for this research whereas unstructured interview would have provided in-depth understanding but likely not relevant information if the interview would not have followed the research topic. All of the interviews were conducted individually with only one interviewee in each interview. This allowed the interviewees to describe their own experiences without being influenced by the views of other participants.

The appropriate sample size required for a qualitative research is not clearly determined because it varies depending on the research. Most importantly, the sample size should enable forming convincing conclusions from the data and thus, it should be defined according to the purpose of the research. (Bryman & Bell 2015, 436.) The term saturation is often used for defining the sufficient amount of empirical data. Regarding interviews this means that the researcher conducts interviews for as long as a new interview provides new relevant information to the research objective. When a new interview repeats the information in the previous interviews, the empirical data has been saturated. (Hirsjärvi et al. 2009, 174–177.) In this research, new interviews were conducted until the information started to repeat itself and the data was found sufficient to form an understanding of the case company’s current state. At that point, the interview data was found to be saturated and the interviews were stopped.

In qualitative research, purposive sampling is typically used instead of probability sampling. Purposive sampling denotes that the research participants are not selected randomly but rather strategically to ensure that the chosen participants are relevant for the purpose of the research. (Bryman & Bell 2015, 428–429.) In this research, the chosen sample included employees from the organization’s financial accounting and management accounting functions. The choice was made based on their role as the major stakeholders of financial data. In addition, data quality has a significant impact on their day-to-day activities and they often need to assess the quality of the data they are using. The interviewees were deliberately chosen to represent data users in different roles, such as data creators and analyzers for them to understand the context and the topic of the research. Including both financial accounting and management accounting employees enabled covering both of their perspectives to financial data.

The interviews were conducted as video meetings because arranging face-to-face meetings was not possible due to the Covid-19 pandemic. To ensure the quality of the empirical data, the interviews were recorded with the permission of the interviewees. This ensured that the content of the empirical data reflected the actual interviews and bias was minimized. The recording also allowed the researcher to give her full attention to listening to the interviewee instead of simultaneously taking notes. When all interviews had been conducted and recorded, the recordings were transcribed to written form word-by-word. In the transcripts, in addition to what was said, it is important to consider how the interviewee said it (Saunders, Lewis & Thornhill 2009, 485), which was included in the transcriptions of this research.

The general strategy for analyzing qualitative data is analytic induction which refers to the use of empirical data to build a theory instead of to test a theory. Inductive research requires a dialogue between the theory and the data to draw inferences out of observations. Hence, data analysis may be started before all data has been collected and the observations may further guide the data collection process. (Bryman & Bell 2015, 579–581.) The data analysis method used in this research was content analysis, in which the data is reorganized in order to achieve a clear and coherent picture of the research phenomenon. In inductive content analysis the data is first reduced to exclude irrelevant data from the data set. Second, relevant data with similar characteristics are clustered to categories, which form the basis to the research structure. Finally, the clustered data is conceptualized to draw conclusions of the data. (Tuomi & Sarajärvi 2018.) Saunders et al. (2009, 481) liken the process of qualitative data analysis to a jigsaw puzzle where separate pieces are grouped together according to similarities to create a clear picture of the empirical data. Inductive data analysis was found suitable for analyzing the interview data in this research because the aim was to formulate a clear picture of the current challenges regarding data quality in the case company. In addition, as semi-structured interviews do not strictly follow the interview structure, inductive content analysis enabled clustering the identified themes together.

The data analysis process in this research followed the process of inductive content analysis by Tuomi and Sarajärvi (2018). The data analysis was started by reading each transcript carefully through. Then, the transcripts were read again by focusing on identifying the challenges for data quality in the case company and listing them on paper. When the challenges had been listed, they were organized to themes according to similarities of the factors affecting data quality. Tuomi and Sarajärvi (2018) note that inductive content analysis does not follow a certain theory. In this research, the data quality barrier themes were used as a high-level classification for organizing the empirical data. Finally, five themes of data quality challenges were identified from the interview data.

3.4 Reliability and validity

When the quality of a research is assessed, the focus is not merely on the methodological choices but rather on the research process as a whole. Reliability and validity are traditionally used as the criteria to evaluate the quality of a research. Reliability in this context refers to stability, consistency, predictability and accuracy. Whereas validity is used to assess how legitimate, correct, justified and watertight the information is. (Kihn & Ihantola 2008, 82.)

In a qualitative research, reliability can be assessed from three perspectives, which include credibility, replicability and systematic evaluation of error sources. Credibility refers to the degree to which the arguments are based on the empirical data, arguments between theory and observations are logical and several data collection methods are used. (Kihn & Ihantola 2008, 92.) In this research, credibility was increased by aiming to make conclusions based on the theoretical framework which provided the basis to understand the empirical data. In this research the empirical data was collected using only semi-structured interviews and triangulation was not used which partially decreases the credibility. Bryman and Bell (2015, 400) add that internal reliability can be increased by including more than one researcher in the research process. This would reduce the individual interpretations of each researcher and provide more objective observations. The interpretations in this research are based on only one person's perception of the observations, which should be considered when assessing the research findings.

Replicability, in other words, the degree to which the study can be replicated is referred to as external reliability by Bryman and Bell (2015, 400). In qualitative research this is challenging to achieve because qualitative research is normally conducted in a real-life social setting, which inevitably alters over time. Therefore, the degree of replicability in a qualitative study is assessed as whether another researcher could replicate the study and would end up to the same results. Most importantly, the research process should be documented in the way that the reader is able to assess the generation and interpretation of observations. (Kihn & Ihantola 2008, 91.) In this research, external reliability was increased by aiming to provide a detailed description of the research process and the collection process of the empirical data to enable the reader to assess the process and the formulation of the findings and conclusions.

Possible error sources in a research process include i.a. poor choice of theory, poor inductive analysis from the empirical data, weak choice of data collection method, ambiguous interview questions and inaccurate transcribing, which compromise the accuracy of the research (Koskinen, Alasuutari & Peltonen 2005, 262–263). These errors in a research can be prevented by conducting the empirical research process carefully and systematically (Kihn & Ihantola 2008, 92). In this research, error sources were diminished by paying close attention to the choices made in the research process. Data collection methods were carefully chosen to fit to the purpose of the research and inductive analysis of empirical data was conducted systematically. The transcriptions were made word-by-word and they were gone through shortly after the interviews to avoid misinterpretations.

Validity in a qualitative research can also be divided to three perspectives, which are internal validity, structural validity and external validity. Internal validity refers to a detailed description of the relevant research phases, making clear connections between interpretations and the data, as well as detailed documentation and logicality of the research process. (Kihn & Ihantola 2008, 88, 92.) Internal validity assesses whether the interpretations are logically made based on the empirical data (Bryman & Bell 2015, 400.) In this research, internal validity was increased by a comprehensive analysis of the empirical data. In addition, an attempt was made to tie the interpretations made from the empirical data to the existing literature to ensure that conclusions were logically made.

External validity refers to the degree to which the interpretations can be generalized to other situations (Kihn & Ihantola 2008, 89; Bryman & Bell 2015, 400). However, in case studies, descriptions regarding people are unique and two similar situations do not exist (Hirsjärvi et al. 2009, 227). Because this research was conducted in a single case company, the data and interpretations are made based on only one case due to which the observations cannot be transferred to other companies or to other situations.

Structural validity in a research is increased by ensuring the transparency and credibility of the formed interpretations and conclusions. In addition, careful evaluation of the fit of the methodological and methodical choices with the research purpose is important for structural validity. (Kihn & Ihantola 2008, 89) In this research, methodical and methodological choices were made based on the goal of the research. The choices were argued based on the literature and described to increase the transparency of the choices. In addition, the conclusions made in the research were made based on the literature to increase the credibility of the interpretations. In addition, citations from the interviewees were used in the description of the findings to increase the credibility of the researcher’s interpretations.

4 FINDINGS

In the fourth chapter the findings based on the interview data are presented. First, the empirical data set is described to outline the details of the interviews. Second, the current state regarding the quality of financial data and processes around data in the case company are described and analyzed. Then, the development of the data governance framework for the case company is presented.

4.1 Description of empirical data

In this section, the data set of the empirical research will be described. The empirical data consists of interviews with the case organization's employees. The interviews followed a semi-structured approach and therefore, they did not strictly adhere to the planned structure of the interview questions. A total of eight interviews were conducted, of which four interviewees were employees from financial accounting team and four employees from management accounting team. The interviewees were chosen to represent the most important stakeholders of financial data in the company meaning that the quality of financial data has the greatest impact on their work. The interviewees from financial accounting department included the financial controller, junior financial controller, payables controller and credit controller. The interviewees from management accounting department included business partner with product, business partner with sales, business analyst and the manager of the team. The interview questions consisted of five background questions and seventeen questions about financial data quality and data governance. Due to the semi-structured nature of the interviews, additional questions were asked outside of the planned question structure if it was found necessary to support the data collection. The interview questions are listed in the appendix 1.

Due to the interviewees' preference, the interview data was collected anonymously and therefore, the names or the titles of the interviewees are not connected to the data presented in this research paper. The interviews were conducted during January and

February in 2021. The total amount of recording was 444 minutes with the average of 56 minutes per interview rounded to the nearest minute. Seven interviews were conducted in Finnish and one in English. The details of the interviews have been listed in the table 3 with the coding used for each interviewee in this research paper.

Table 3. Details of the research interviews

Date	*Function*	*Duration*	*Coding in the research*
21.1.2021	Financial accounting	56 min	FA1
29.1.2021	Management accounting	43 min	MA1
2.2.2021	Financial accounting	57 min	FA2
4.2.2021	Management accounting	61 min	MA2
5.2.2021	Financial accounting	66 min	FA3
8.2.2021	Management accounting	51 min	MA3
9.2.2021	Financial accounting	65 min	FA4
10.2.2021	Management accounting	45 min	MA4
Total of 444 min	8 interviews	An average of 56 min	

4.2 Current state in the case company

This section describes the current state of the case company regarding data quality and data governance. The identified challenges barriers to the quality of financial data are described and the current processes around financial data are clarified. This section seeks to find the answer from empirical data to the first research question: What kind of challenges exist regarding financial data quality?

The interviewees were asked to describe the challenges they had identified regarding data quality in financial data. The identified challenges were classified according to the three themes found from literature, which included management, roles and responsibilities, communication, and technology related challenges. In addition, a new theme was identified, which was named internal conditions.

Management

The current challenges related to management were found to be the lack of leadership and overview of financial data and the lack of structure in data processes. Because coordination was missing, it was not systematically ensured that all data processes work effectively together. In the case company financial data was generated and used in several operational processes by several people but there was not a clear view of financial data as a whole. Due to this, the big picture was not always taken into account when the focus was on the operational processes and thus, the big picture had occasionally suffered. Hence, financial data as a whole had not been actively managed.

> *Nobody has probably taken the responsibility in a way, that hey, let's do this thing and really look into it. Now we've realized that we need to reconsider our processes and assess whether everything is working as it should. ... At this point we've become quite a big company so I think it cannot be underestimated that leadership in these things is important. (FA4)*

> *What has been missing is that we state the things we want to follow, do it and then at some point later take a look whether something needs to be updated. But it has been like the Framery style that changes are made on the go and the big picture has suffered in some cases. (MA4)*

The interviewees were asked to describe how the financial data activities were currently organized in the company. All of the interviewees agreed that a formal structure for data activities did not exist. However, five of the interviewees described that there was some kind of consensus of way the processes were organized, even though it had not been intentionally formed or documented. The informal way had been adequate for conducting the operational data activities so far because the operational things were running even though the structure was lacking.

> *I wouldn't say it's organized but it has formed to its current shape accidentally in a way. We do things in a certain way even though it's not based on anything defined. It's just that this is how it's been done. (MA1)*

Processes where financial data was generated and used were not defined or documented in the case company, which occasionally caused unclarity of how they should be conducted. Regarding operational data activities which were conducted on a daily basis, the processes were clear to the employees who conducted them. Data processes, which were conducted only occasionally were perceived as quite unclear, such as making changes to the master data objects.

> *Within the finance team we have quite clear tasks and processes from which the financial data is generated. But in case of change requirements, then I don't know if a clear process really exists in that. (FA3)*

There was not a clear process for conducting the changes to master data, and thus, they often happened in ad-hoc style. In addition, there was typically a tight schedule for conducting the needed change, which required quick action and decisions (MA3).

> *How I see it (changes to master data) is that a need is identified somewhere and then it's communicated in Slack to some group channel and then the decisions are made case-by-case. I think there's not a clear process or at least it's not so clear that everyone knows exactly how it goes. (MA3)*

Among the interviewees, a clear part of the process of a master data change was that the IT team was responsible for making the changes to the accounting system. The open questions about the process regarded who makes the decision for approving the change to the master data, who decides how it is named and how is that communicated to the relevant stakeholders that the change affects (MA4).

Undefined processes also hindered the awareness and understanding of financial data processes as a whole. Financial data was used for several different purposes and by different people and thus, without defined and documented processes, the relationships between the different data processes remained unclear. This created a risk that when decisions regarding financial data were made, all the processes that would be affected by the decision were not taken into consideration because there was not adequate awareness of the effects that the change would have further in the data chain. As an example, MA2 had encountered problems in reporting when a change had been made to the generation of the data, that his main reports were based on. The decision had been made elsewhere in the company and he had not been aware of the coming change, which then created problems for the reporting. Even though he was a key user of the said data, the effect of the change on his reports had not been understood beforehand because the relationships of the data and data processes were unclear. Therefore, documenting all processes related to financial data and their interrelatedness was essential to achieve adequate understanding of them. This would enable taking all necessary stakeholders into account when a change was planned to make sure that it would be beneficial for everyone's work. (MA4)

Understanding of the big picture of financial data would have been necessary to have a proactive understanding of the company's data needs, which was mentioned by two interviewees (FA4, MA4). FA4 had encountered a situation where she had to work with data, of which collection process had not been thought beforehand. Thus, she had to manually collect the data from the system which was a laborious task. This was due to the fact that the reporting needs for data had not been considered before the reporting process was started.

> *When I started to report the figures from 2019 … I noticed that it had not been very clear or that it had not been thought beforehand that these figures should be reported to somewhere at some point. And then it was a bit of a patchwork to export the data to Excel. It was purely manual patchwork to build a balance sheet for a certain company … It had not been taken into consideration of how the data will be used in the future. (FA4)*

MA4 noted that proactive understanding of data is needed to ensure that changes in the company are timely reflected on the data to maintain high data quality.

> *For example, regarding departments, if some organizational changes are made, we immediately consider how that affects our department logic and whether some changes need to be done, create a new department, or transfer employees between departments. Proactively make the changes to the systems and sync them with the HR systems. It has been very reactive so far but there are possibilities to make it more proactive when we understand the impacts on the things and departments. And also, the people in the departments understand to communicate certain organizational changes to our business partners on time. (MA4)*

Roles and responsibilities

All of the interviewees described unclarity regarding the responsibilities and decision-making authority regarding financial data. This was because the roles and responsibilities regarding financial data had not been defined in the company and thus, there was not a clear consensus about them. In the earlier years of the company, the finance team was still quite small and the decisions regarding financial master data were made by the CFO (FA3). Since then, the company had grown and employees in the finance department had increased which had blurred the decision-making authority. (MA4)

The interviewees described uncertainty of the situations in which they could make a decision by themselves and in which someone else should be asked (FA2). It was also unclear from whom a certain decision should be asked. For example, if a change was needed to a master data object, such as adding a new department, it was unclear who would make the final decision about it. (FA3) So far in these situations, the change had been discussed within the relevant team and the decision had been made together. However, this created a risk that all stakeholders were not taken into account in the decisions because the big picture was not clear. Thus, the current teams were too big for collective decision-making. Again, operational activities were working out well but the lack of defined decision-making authority created a risk that the control over available data would be lost if decisions were made by several people. This would eventually be reflected on the data quality in inaccuracy or inconsistency problems. At this point of the company's history, it was necessary to have clearly defined responsibility areas for the data to ensure that all relevant aspects were considered in decision-making.

As employees had not been dedicated responsible for certain data, the responsibility was spread among several people. Due to collective responsibility, the responsibilities were not clear to the individual employees. On one hand it created a possibility that several people were conducting the same tasks because it had not been clearly agreed whose responsibility the task was. On the other hand, it also made it possible that nobody took a responsibility over a certain task because everyone thought that someone else would do it.

> *The company has never really taken time before to write down processes because processes were changing all the time. ... So we have never really had clear ownership of things and processes and these kind of things. There have been some cases where people might be doing double work and so on… So that's something that has happened and I think will happen until we get there (MA1)*

> *I think that some processes flow quite smoothly and in some there's a bit more unclarity of who's taking the responsibility. And then what I've noticed is that if nobody takes the responsibility it may be totally left undone. As everyone seems to*

> *have their hands full and if a clear responsibility is not assigned, it may be found quite vague. (FA4)*

The lack of clearly defined roles and responsibilities also hindered the employees understanding of what is expected of them regarding financial data. For example, certain data dimensions were added on vendor invoices to classify expenses to different categories for management accounting purposes. However, it was not always clear to the employees whose responsibility it was to add these dimensions on the vendor invoices (MA1). Thus, the invoices had to be corrected afterwards or did not provide reliable information for management accounting. In addition, it was found unclear who is responsible for the customer master data as several teams were involved in using it (FA1, FA2).

> *There are a lot of grey areas and uncertainty. If you think about customer data, it comes from the CO-team but there are also sales representatives and different systems and so on. The daily work runs okay but to my understanding there are no clear roles or responsibilities. (FA2)*

Communication

All of the interviewees described challenges related to communication that affected data quality. The main challenges identified were the lack of communication, unawareness of data quality requirements and unclear definitions for data. The interviewees found that within teams, communication was on a good level but the lack of communication was noticed regarding communication between different teams (MA4). Thus, data quality issues were often discussed within teams but not communicated to other teams that the issue potentially concerned. The lack of communication between different teams was also found to be one reason for unclarity of processes and understanding the big picture of financial data. Awareness of how data is created and used in different teams was found important to increase understanding of financial data as a whole (MA4).

Communication about changes that had been made to the data to the necessary stakeholders was occasionally found insufficient according to the interviewees (MA1, MA3). When a change had been made to a process or the data itself, it was important that all employees whose work would be affected by it were informed about the change.

However, this was not always communicated on time to the correct people, which had occasionally led to problems. The lack of communication about changes in the data was partially due to the lack of awareness of how the change would affect the tasks of other data users. In addition, a clear responsibility for who should communicate the changes to data users had not been defined.

> *If a certain practice is changed, for example to which field certain information goes, it's often not communicated widely enough. It's logical that it's communicated to people, who's operative work is affected by it. But that the information might affect other people's work afterwards through the data. That's something that's not viewed broadly enough that who are the users of the information in the end. (MA1)*

Communicating the changes to all necessary data users was seen as essential to retain the data consistency.

> *And now the need from management accounting is that the number of projects is increased and departments are more clearly separated. That requires the right kind of communication at the right time to the right people. To make all the necessary people aware that what needs to be added and where, if something needs to be added. (FA4)*

The lack of communication hindered data creators' awareness of the data quality requirements that data users had for the data. If the data was created in a different team than it was used in, communication was essential to create awareness of the data quality requirements to the data creators. Without the knowledge of the requirements, the data would inevitably include defects. Due to the lack of communication, data creators often were not aware of how the data was used later in the process. This hindered the understanding of the data process which decreased the data creator's ability to ensure the accuracy of the data (FA2).

> *Probably the biggest reason is that it's not understood why certain project costs are being followed. Necessarily there's not even awareness that such project is available for use and there's not understanding of what the data is used for. It might seem a very small thing for one person but be highly important for someone else's work. So, the reason is lack of understanding, lack of communication. (MA4)*

Communication of the data quality requirements for different employees involved in the data process was seen as one of the most important factors that affected data quality.

If corrections are always made in behalf of other people, it's not very efficient in the long term. When you tell people, what has gone wrong and what needs to be corrected, you'll soon notice that it gets better and more effective and the attention is put on the right things. It's not anyone's fault but I think it's just been the lack of communication about these things. (FA4)

Communicating data quality requirements was seen as especially important in cases when financial data was created outside of the finance team. This was because the external regulations set certain requirements for financial data but they were not clear to employees without financial background. As an example, customer operations team in the company is in charge of creating the sales invoices in the accounting system. Sales invoices must comply to several regulations, which determine i.a. the date and the tax rate to be used on the invoice. Thus, the employees creating the invoices must be made aware of the requirements with clear communication and instructions. Also, purchase orders that affect vendor invoice processing are created in the sourcing team. Purchase orders require information of the correct bookkeeping account, but the correct account is not always clear to the purchase order creators without accounting background. (FA2)

Quite few people for example understand accounting as such. It's seen as a distant thing. Even when someone might daily do things that affect accounting. ... An example are sales invoices. Sales invoices are created but it isn't necessarily clear where it eventually affects and how important it is that dates are correct and on a correct period. (FA4)

The main thing is that the accounting understanding is limited for people who have had little to do with it. For example, they don't know the accounts of the profit and loss statement. When they order something, they don't know whether it goes to purchases or supplies. ... If the account on the purchase order is incorrect, it's transferred to the vendor bill and I have to manually correct it. (FA3)

Three interviewees (FA3, MA3, FA4) described improvements in data quality that they already had noticed when the data requirements had been discussed with the data creators. This highlights the impact of communication for achieving high data quality.

The effectiveness of communication was found to be negatively affected by unclear definitions for data. MA4 described that certain terms were not understood and used similarly among all employees. Thus, when terms were understood differently in discussions, it had sometimes led to misunderstandings. The lack of common understanding decreased the awareness of data requirements and hence, affected the data quality. Because employees from all departments in the company operated with financial data, it was essential that things were communicated in the way that everyone would have common understanding of financial data regardless of their background. Common understanding of data was seen as key factor for effective communication and achieving high data quality. (MA4)

Technology

Technology related barriers included factors in the information system or other tools that hindered achieving high data quality. The identified challenges were manual tasks and information system characteristics. Four interviewees (FA1, MA2, FA2, FA4) mentioned the negative effect that manual tasks have on data quality. Manual tasks were found to increase the possibility of defects due to human mistakes as typing errors or careless actions in data entry could lead to errors in data. Even though manual tasks regard the information system aspect, FA3 noted that it is also an issue related to instructions and the data creator's motivation.

> *We still have quite a lot of manual tasks which I find to increase the number of human mistakes. When things are done in a hurry, something small can occur to anyone. (F4)*

Automation was seen as an important way to reduce manual tasks and defects in the data. (FA2, FA4) However, the interviewee F4 also noted that the automation of the systems cannot be blindly trusted. This was because she had observed cases where the automated process of certain data creation had been affected by a system update and the automation had created incorrect data after the update. Thus, even automated data creation requires some kind of monitoring to ensure its quality.

The employed systems also occasionally caused defects in data. As an example, sometimes the system was found unable to collect certain data. Even though there was a need to include a certain dimension to data, occasionally it was not possible due to the system characteristics (FA4). In addition, customer data was generated in a system outside of ERP and the data was transferred to the ERP through integration. Due to the different data fields in the systems, the customer data transferred to ERP was incomplete.

> *The problem is that it's not integrated yet, so typically when sales managers create an account, they cannot add the customer industry. Even though that would be the best because you know they have the best knowledge of the customers, they know their industry so they should actually be entering that information. So, it doesn't kind of flow to the ERP. (MA2)*

Internal conditions

The identified challenges that were classified as related to internal conditions regarded the internal factors in the company that affected data quality and the processes. The challenges related to internal conditions included scarce resources, changes in daily operations and frequently changing data requirements which were mostly due to the company's fast growth in the past few years.

These factors related to the fast growth in the company during the past years were mentioned by seven interviewees as reasons for data quality challenges. When the company's processes were trying to keep up with growing sales, the resources were scarce and things had to be prioritized. This meant that everything could not be done in high-quality manner. The operational activities were successfully conducted but the processes did not evolve along with the business.

> *It's because not that long ago the company's turnover was much smaller, much less pods were shipped and sent and the day-to-day operations then were focusing on getting the pods out there rather than really measuring like how much does it cost to actually make a pod. But now that we are getting bigger, we need to be able to have that kind of information so it's also about kind of the company's history and resources. (MA2)*

Regarding the data used in financial accounting, it had always been essential that it complied to the regulations and it was correctly reported to external stakeholders, which is why it was prioritized and its quality had stayed on a good level.

> *I think the reason is that the speed of growth has been so exceptional. When the people and the system are trying to keep up, I find it very humane that data management issues are coming a bit later. I would be surprised if everything was*

> *managed to keep perfect all the time. ... I believe the speed has been high and the minimum has been that everything is working and we can do business. And to be able to report the right figures to external stakeholders, such as tax authorities and so on. That has been the minimum and now we are making the improvements. (FA2)*

The scarce resources was seen as a challenge to data quality because without sufficient resources, data quality was not paid enough attention to. One interviewee (FA4) noted that the lack of resources had been visible in the situation that certain people had too many responsibilities in terms of their resources. Due to this, things often were not planned ahead and all necessary aspects of data were not considered when things were conducted in a hurry. In addition, employees that were caught up in the operational tasks lost the bigger picture.

> *In a company like Framery when the speed is high, we might not know to stop at the right places. (MA4)*

FA4 had also noticed a clear improvement in the quality of data when the resources for considering data quality had grown, which implies that adequate resources play an important role in achieving high quality of financial data. Also, M4 noted that data quality had been recently put in the focus since the resources in financial accounting and management accounting teams had been increased.

In the early stage of the growth, the company was still small and between just a few people the coordination was effective without formal processes. It was even beneficial to keep the hierarchy as low as possible. As one interviewee put it:

> *When there are only a few people, there cannot be very strict restrictions for who can do what to keep things working. ... Rather, it was beneficial that everyone was able do everything. But it's not anymore. (FA2)*

When the company grew, constant changes had to be made in the organization and its operations. The changes created challenges to data quality as new processes were often conducted in ad-hoc style. Also, because the processes were constantly changing, the company had not taken time to write them down (MA2). New recruitments were made frequently which meant that the teams were constantly growing and changes were made to the team structures, which also affected the data.

> *When the pace of growth was at its peak, the teams were changing, everything was changing and new things were coming, the internal communication and the lack of it surely led to problems in the data quality. ... Which was due to the fact that the situation was not stable in any way. (FA3)*

Due to the fast increase of employees, the coordination spread out to several people. When the coordination was spread, and the processes were not defined, this created risks to data consistency. Now as the company had grown, the interviewees (MA4) agreed that the processes are needed to bring clarity and certainty to working with financial data.

> *The amount of people has grown here so much that surely the ball has dropped a few times along the way. (FA4)*

Four interviewees (FA2, FA3, FA4, MA4) mentioned that the data quality was affected by the frequently changing requirements for the data. It was common that instructions and requirements for certain data were changed frequently as new information was acquired, which influenced the data quality (FA3). Additionally, several needs formed on the go, which required fast changes to data (FA4). When the requirements change fast, it creates challenges to the data accuracy because it requires that the data creators have always to most recent knowledge about the data requirements. The changing requirements also compromise the consistency of data, when new changes require that things are done differently than in the past and thus, the comparability of the data is affected.

> *For example, regarding the subsidiaries, the instructions of things, such as in which countries we need to be VAT-registered and in which not. The instructions have been changed and redefined which affects how the invoice is handled and then also the data. (FA3)*

> *I started to write a tax manual a year ago and I still open it every week and make some changes to it. In a way it will never be ready. (FA4)*

Assessment of data quality

Despite the described challenges, the interviewees assessed the overall level of the financial data quality as quite good and being in a stage of development. Additionally, data defects were often searched for and corrected if incompleteness or inaccuracy problems were noticed. In terms of financial accounting tasks, the data was assessed as being on a good level, whereas the data used in management accounting was described as being in a state of development.

> *As we want to move on to the direction that we want to get more out of the data in terms of management accounting, I see that being in a development state. But overall if I think about it (data quality) from the perspective of my own work, it's good. I don't recognize any big defects. (FA2)*

This reflects the history of the company. Even when the growth of the company was fast and resources were scarce, complying to external regulations had to be prioritized in the company. Whereas management accounting data does not follow legal requirements and thus, the resources were put on other tasks. In addition, the needs for management accounting were not very high in the beginning.

> *When the numbers are small, they can be quite easily analyzed and if you start to build a lot of management accounting calculations on top of it, you don't get the same value from it. When the mass is bigger, then there's actually a need to put it into smaller pieces (for analysis). (FA2)*

The current management accounting team started to grow on the spring 2020 and thus, after recognizing the needs, the requirements for data quality are still being formed. In conclusion, the quality of the financial data was found to be on a good level but the need for development was seen in the processes to ensure maintaining high data quality in the future.

Consequences from poor quality data

The interviewees also mentioned consequences that insufficient data quality had on their work. The consequence that was mentioned by six of the interviewees was the additional manual work from validating and correcting inaccurate data. Data defects were corrected to prevent their effects on reporting and analysis but it was found time-consuming. The interviewees saw that the time used on manual corrections decreased the time that could be spent on value-adding activities.

The efficiency of processes was also seen to be negatively impacted by poor quality data. For example, if the email address of a customer was outdated, the collection activities lost effectiveness as the collection activities could not reach the customer. In addition, data quality problems in customer master data were seen to potentially have a negative effect on customer satisfaction.

Especially the four interviewees from management accounting mentioned data quality to have a major impact on decision making. This is because the trustworthiness of reports was found to be highly depending on the quality of the underlying data. If the data was perceived as unreliable, it could not be used in decision making. If there was not awareness of flaws in data and the information was used, it could be misleading and result in wrong decisions.

4.3 Needs for improvement and expected benefits

The interviewees were asked to clarify, what kind of improvements would they want to see regarding financial data processes. These needs were mostly in line among all interviewees and reflected the identified challenges.

Firstly, six of the interviewees (FA1, FA2, MA2, MA3, FA4, MA4) mentioned that ownership and responsibilities for financial data should be determined and clarified. Defining data ownership referred to decision-making authority and accountability for certain data to clarify who makes decisions regarding data, such as making changes to master data. Equally important was found to define responsibilities in the data processes. Responsibility areas were needed to be defined to ensure that it was always clear to the employees who is responsible for which task. In this vein, clearly defined responsibilities would increase employees' understanding of what is required from them regarding financial data.

Secondly, five of the interviewees (FA2, MA2, FA3, MA3, MA4) found necessary, that the data processes were defined and clarified. Defining processes was found important to clearly see the relationships of the different data processes. Making that transparent would improve understanding of how a change in one place affects other people's work further in the data process (MA3). In addition, especially the process of making changes to master data needed clarifying according to the interviewees. Because new needs to the master data would continue emerging, it was important that all changes were made according to the same process to ensure the consistency of the data. One interviewee (MA3) importantly noted that processes must be at place to ensure consistency but they should not make the work too stiff. Flexibility in the processes is needed for the company to be able to quickly respond to changes.

Thirdly, five of the interviewees (FA1, MA2, MA3, FA4, MA4) mentioned needs regarding increasing communication between different teams. Communication was needed to facilitate employees' understanding of the data processes and hence, the awareness of the requirements for data quality. In turn, understanding the data processes would increase information sharing, when all necessary stakeholders were known. Better instructions were also mentioned necessary to decrease the possibility of human error in complex tasks (FA4). In addition, it was found necessary that terms were better defined and shared company-wide to facilitate common understanding of them and thus, contribute to effective communication (MA4).

Fourthly, one interviewee (MA1) found important that all available financial data was mapped and defined. After the mapping, the purpose of the data should be determined and only relevant data should be retained in the system. This requires understanding and determining what is important data for the company to follow and what is not. Mapping the data would also assist in clarifying the relationships and uses of different data.

The interviewees described three kinds of benefits that they would expect from better governed financial data. Seven of the interviewees mentioned that through better data quality the time that was currently used to manually validate and correct the defects in the data would decrease. This time could then be used on value-adding activities such as analyzing exceptions in the data. Especially interviewees from financial accounting mentioned that high-quality data would lead to increased trust on figures on financial reports as well as ensure compliance. All four interviewees from management accounting emphasized that through better quality data, the created reports would be more reliable and provide more valuable information for decision making. Hence, the decisions made based on the data would be more accurate and benefit the company. One interviewee also noted the importance of high-quality data on operational efficiency. Debt collection process would not work effectively without high quality data and in the worst case could impair the customer relationships. This highlights the importance of well governed data and high data quality for the company.

4.4 Summary of the key findings

Management

The main data quality challenges related to management were the lack of leadership and overview of financial data as well as the lack of structure in data processes. Because the overview was missing, financial data as a whole was not taken into consideration in decision-making. In addition, without the consideration of the big picture financial data, changes in the organization were reactively reflected on the data instead of proactively ensuring the quality of data. Regarding financial data processes, the interviewees agreed that the process had not been defined but the operational work was not currently significantly impacted by that. Undefined processes occasionally created unclarity in tasks and resulted in ad-hoc actions especially regarding changes being made to master data, which was seen to pose a risk to data consistency. In addition, the lack of defined processes hindered the understanding of financial data as a whole. Regarding management related challenges, a need for defining all financial data and data processes were mentioned to gain understanding of the big picture of financial data and data processes.

Roles and responsibilities

Roles and responsibilities had not been defined in the case company, which lead to unclarity in decision making and tasks regarding data. Undefined decision-making authority often led to collective decision making, which created a risk that all stakeholders were not considered. In addition, individually made decisions created a risk for data inconsistency as the overview of the data was missing. Because responsibilities regarding financial data had not been defined or documented it was not always clear to the data creators what is expected of them regarding financial data. Occasionally the employees had made double work but also tasks were possibly left undone if nobody saw it as their responsibility. Thus, undefined responsibilities compromised data quality as accountability for the quality of certain data had not been allocated. Defining data ownership and responsibilities regarding financial data was mentioned as a need for improvement by the interviewees to clarify the responsibilities of each data stakeholder.

Communication

The main challenges related to communication included the lack of communication between teams, unawareness of data quality requirements and ineffective communication due to misunderstandings. Due to the lack of communication between teams, data quality issues were often discussed within teams, which hindered the understanding of data processes around the organization. Communication was also seen to affect data creators' awareness of data quality requirements. Especially as the requirements of financial data are externally regulated, without communicating them to the data creators, it could lead to defects in data. Due to the lack of communication, the data creators often were not aware of how data is used later in the process, which was found to limit their ability to ensure the quality of the data. In addition, due to unclear data definitions, there was not common understanding of all terms related to data in the company which was found to lead to misunderstandings. When misunderstandings occurred, the communication was not effective and potentially led to defects in data. Effective communication between teams was seen as essential for achieving high quality of financial data. Therefore, the interviewees described a need to increase communication of data quality issues between teams and establish common definitions for terms to facilitate common understanding of financial data in the company.

Technology

Technology related challenges included manual tasks and information system characteristics. Manual tasks were seen to increase the probability of data quality defects due to human mistakes. Automation was seen as a possible solution for this, but one interviewee had observed that data created through automation had occasionally also included defects, so it could not be treated as a panacea. In addition, the system had sometimes found unable to fulfill certain data quality requirements because of the system characteristics. Also, integrations between different systems were observed to cause defects to data completeness.

Internal conditions

Challenges that were categorized under internal conditions included scarce resources, changes in daily operations and frequently changing data requirements. These are factors that are due to the internal conditions in the company that cause challenges for achieving high data quality. Because resources in the growing company were scarce, things were often not planned ahead data quality was not paid enough attention to. The interviewees mentioned that recently increased resources had enabled better consideration for data quality. When operations in the company changed, it consequently led to changes in data. When several changes were made to the data without clear coordination, this created a risk for data quality, if all data stakeholders were not aware of the changes. Frequent changes to data requirements created a challenge to data consistency. When data quality requirements were often redefined, achieving data consistency was challenging as the comparability of the data was affected. In addition, frequent changes required that the data creators have always accurate knowledge about the data requirements.

4.5 Developing a data governance framework for the case company

In this section, a proposal for a data governance framework for the case company is presented. First, the roles and responsibilities to the case company were formed on the basis of the model by Weber et al. (2009) and Cheong and Chang (2007, 1005). Then, data governance activities were designed according to the data governance domains in the model by Khatri and Brown (2010). The aim of the framework is to provide the necessary governance for financial data processes to ensure high data quality and provide clarity for the people working with the data. In addition, the framework aims to fulfill the improvement needs from the interviewees that were presented in the previous section.

Roles and responsibilities

The first step in developing data governance framework is establishing clear roles and responsibilities for data. This is because the roles may then be assigned as responsible for the data governance activities. (Alhassan et al. 2019, 106.) Therefore, the structure of roles and responsibilities for data governance framework was designed first. Roles and responsibilities identified in the literature were introduced in the theoretical framework section 2.2.2. Weber et al. (2009) state that data governance framework should always be developed to each company individually to reflect the company's unique contingencies. Thus, also the roles and responsibilities are defined in each company individually by including the roles and decision areas that are relevant for the company's needs. (Weber et al. 2009, 9.)

Weber et al. (2009, 14) use a continuum between centralized and decentralized to describe the organizational structuring of data governance activities. Centralized data governance approach denotes that the chief steward or the data quality board has the decision-making authority. Whereas in decentralized data governance approach the decision-making authority is placed on data stewards. Companies should balance their structure between these two ends according to their contingency factors. Weber et al. (2009, 19) describe that centralized approach fits companies in which strategy is profit-focused, firm size is small and organization structure is centralized. Decentralized approach is suitable for companies which have a growth-focused strategy, firm size is large and organization structure is decentralized. (Weber et al. 2009, 14, 19.) Because the suggested framework only covers the scope of financial data and regulations guide data requirements, it is suggested that the approach leans to centralized decision-making style. This is because achieving high data quality requires that all data objects work effectively together and thus, the overview of financial data is required in decision-making. In addition, because financial data affects compliance, a centralized decision-making could better ensure that individually made decisions do not harm data quality. Decentralized approach was seen beneficial for allocating responsibility for individual data objects where more detailed knowledge is required. Therefore, it is suggested that responsibility and decision-making are divided to different levels.

In the literature, data governance responsibilities are often divided to strategic, tactical and operational levels (Cheong & Chang 2007, 1005; Ladley 2012, 127). These levels define the scope of responsibility of each role. These levels of responsibility were also seen suitable for the case company. This is because low hierarchy was part of the culture in the case company and thus, unnecessary hierarchy in the roles was tried to be avoided. In addition, the data governance framework has to enable flexible changes to the data while simultaneously ensuring data quality and effective operations.

Strategic view over financial data assets is needed to ensure coordination and that the different data processes work effectively together. This is because a challenge identified in the interviews was that nobody was in charge for the overview of financial data and thus, the big picture had occasionally suffered. Because data needs would only increase in the company in the future, it is suggested that strategic responsibility and overview of financial data assets is allocated to one person to ensure effective coordination of data objects. Because of the responsibility of the overview of all financial data, the person should ensure that the decisions benefit financial data as a whole and all necessary stakeholders are considered. A possible role for strategic responsibility could be a data owner. Data owner is described as belonging to a certain business department (Otto 2011a, 242) and being accountable for the data assets in that department (Abraham et al. 2019, 428). In the domain of financial data, the data owner should be knowledgeable of the requirements for financial data quality because she/he defines business requirements for data (Otto 2011a, 242). Therefore, it is suggested that a data owner is recognized in the company. The data owner could then be assigned responsible for the overview of the data assets and data processes as well as ensuring that data assets are developed in a consistent and sustainable way. Cheong and Chang (2007, 1005) place strategic responsibility on data governance council. Because financial data is used in several departments of the organization, a data governance council consisting of different departments using financial data may also be established to balance interests from different departments.

Tactical scope refers to medium-term view over data assets. Cheong and Chang (2007, 1005) assign data steward as responsible for tactical scope of data. Data stewards should have detailed knowledge of data processes and requirements (Cheong & Chang 2007, 1005). Data stewards also support data stakeholders in data use and evaluate problems with data (Otto 2011a, 242). Detailed knowledge of financial data is required especially regarding the use of master data objects. Therefore, it is suggested that tactical responsibility is placed on the data stewards according to individual master data objects. Due to the large amount of master data objects, there should be several data stewards in the company to ensure detailed knowledge of each master data object. Thus, it is suggested that the responsibility of data stewards is to ensure effective and consistent use of individual master data objects. This includes maintaining the data objects and conducting necessary changes to data. In addition, they communicate the data requirements to data stakeholders and guide data users in using the data appropriately. Data stewards communicate with the data owner and the data users to ensure common understanding of data.

Operational scope refers to the day-to-day activities, such as creating, processing and analyzing data. Cheong and Chang (2007, 1005) allocate operational responsibility to user groups who consist of data stakeholders involved in the data activities. Abraham et al. (2019, 429) separate data producers as the role that creates or maintains data and data consumer as the role that uses data. Data users are responsible for reporting data problems, requesting data needs and specifying reporting requirements (Cheong & Chang 2007, 1005). The role of data users was not commonly mentioned in the literature but in the case company it was found important to clarify the roles and associated responsibilities to all people working with financial data. This is because a challenge identified in the interviews was that data producers were not always aware of their responsibilities regarding data. Therefore, specifying the roles would mitigate this problem. Whereas data stewards are responsible for communicating data requirements to data users, data users are responsible for producing and using the data according to these requirements and reporting issues or needs to the data stewards. Figure 4 presents the structure of the suggested data governance roles.

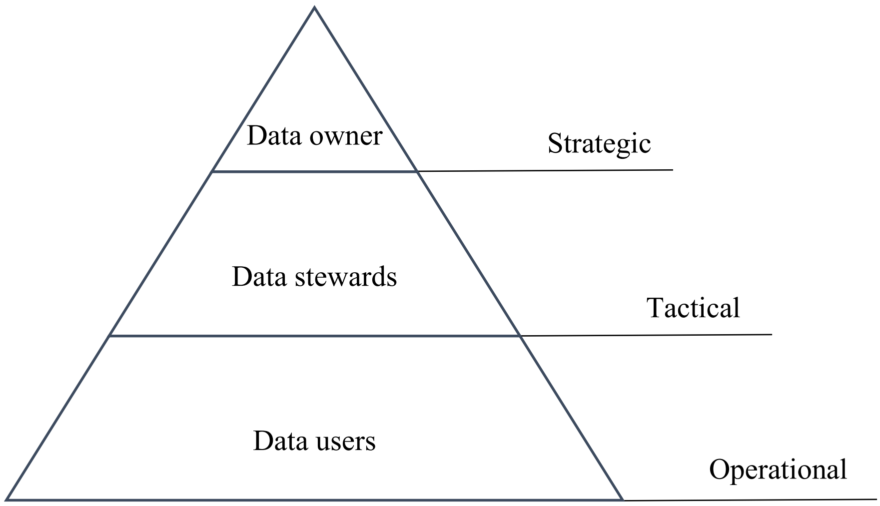

Figure 4. Data governance roles in the case company

Data governance activities were designed in the case company according to the data governance model by Khatri and Brown (2010). The model includes five data governance domains which include data principles, data quality, metadata, data access and data lifecycle (Khatri & Brown 2010). The designed data activities are clarified next.

Data principles

The first decision domain in the model is data principles. The key goal for establishing data principles is to clarify the role of data as an asset in the company. Data principles determine the desired behavior regarding data and thus, form the direction for other data governance decisions (Khatri & Brown 2010, 149–150). Data principles link a company's strategic goals to overall data governance framework by establishing a data strategy (Brous, Janssen & Vilminko-Heikkinen 2016, 120; Alhassan et al. 2019, 102, 107). A data strategy should include the long-term objectives that are pursued with data and how they are aligned with the company's long-term business goals (Brous et al. 2016, 120).

Therefore, it is suggested that a data strategy is established in the company to determine the key goals that are pursued with financial data. Strategic objectives for financial data may include improving decision making with reliable data, ensuring financial compliance and developing operational processes. These goals should then be aligned with the company's strategic goals to ensure that the data strategy supports the business strategy. Documenting the data strategy forms a business case for data governance because it highlights the need to manage financial data as an asset. The strategic goals establish the long-term direction for data activities and data governance should be designed to support these goals. In addition, strategic management of data requires that strategic initiatives are clear. These strategic goals for financial data should be collectively determined within management and data stakeholders by basing them on company goals.

Data strategy is needed to create a company-wide understanding of the importance of financial data. When data is linked to company objectives, the linkage to daily processes is clarified. Also, a need to have a proactive understanding for data was identified as a challenge to data quality. Data strategy provides a strategic view over data and provides the direction for strategic data management. The it is suggested that data owner is responsible for overseeing the data strategy as the data owner is also responsible for the long-term view of financial data.

Data quality

Data quality is the second data governance domain. It refers to establishing the requirements for data quality, which are determined by the context of the use of data. (Khatri & Brown 2010, 149–150.) In the case company, data quality requirements had not been formally defined or documented which resulted in unawareness of the requirements among data stakeholders. To address this challenge, it is suggested that data quality requirements are defined and documented in the case company. The requirements should include information about what data fields are mandatory, what information and in which form should the fields include and who is responsible for entering this information.

External and internal regulations are a key element in defining data quality requirements (Alhassan et al. 2019, 107). This aspect is especially important in the context of financial data because the data requirements are influenced by external regulations, which the data quality requirements must comply. Thus, the quality requirements should be determined by someone who has an appropriate understanding the externally set requirements for data quality. Alhassan et al. (2019, 107) suggests that business owner should initially understand the data requirements and communicate them further. In this data governance framework, it is suggested that the data stewards defined the data quality requirements for the data under their stewardship. This is because they should have detailed knowledge of the data requirements (Cheong & Chang 2007, 1005). Data owner should then be responsible for overseeing the requirements and ensuring that they comply to external requirements. Whereas data steward defines the requirements for data quality, data creators are responsible for creating the data according to these requirements.

A key part of setting data quality requirements is communicating them to data stakeholders (Khatri & Brown 2010, 149). This is because understanding them is a prerequisite for the data stakeholders to be able to fulfill them (Alhassan et al. 2019, 107). Hence, after defining the data quality requirements in the company, they should be communicated to data stakeholders. This ensures that all data stakeholders are aware of the requirements for the data they create or use. In the interviews it was mentioned that there was not a clear responsibility for communicating data requirements to data users and thus, the responsibility should be defined. From the perspective of the data quality roles, it is suggested that data steward is responsible for communicating the data quality requirements to the data users as they operate closely with them and guide them in data use. The data steward should understand how the data is created or used to be able to communicate and advise data stakeholders.

As was mentioned in the internal conditions, requirements for data might change frequently. Because data quality requirements change when the data is changed, the data steward should also be responsible for updating the data quality requirements and communicating them to data users on time. Setting data quality requirements and communicating them increases the awareness of the requirements for data among the data creators which was identified as a challenge in the case company. In addition, communication about the data requirements between data stewards and data users increases the understanding of the data processes between both parties. Ensuring that all data stakeholders are aware of what is expected of them regarding data quality is essential for achieving high data quality. Data quality requirements are established and communicated to address the challenge of lack of awareness of data quality requirements among data stakeholders.

Metadata

The third data governance domain is metadata, which refers to establishing definitions for data so that it is interpreted similarly by all users (Khatri & Brown 2010, 149). Referring to Smith (2007), Brous et al. (2016, 121) note that a perquisite for successful data governance is understanding the meaning of the data and its importance to the company. In the context of the case company, a challenge in communication was unclear terms that were understood differently among data stakeholders. Therefore, it is suggested that the company establishes metadata, in which the financial data is defined and documented in a standardized way. Metadata should be able to connect the data definitions to real-world concepts (Khatri & Brown 2010, 150) and thus, comprise a business data dictionary (Weber et al. 2009, 12) which is especially important in the context of financial data. Because the terms for financial data may be understood differently outside the finance function, the data definitions should translate the finance related terms in an understandable way to other data stakeholders. Metadata may include information about what is the purpose of the data and what roles are associated with it, such as the creator or modifier and it may be specified at different organization levels (Khatri & Brown 2010, 150). Regarding financial data it is suggested that metadata in the company is specified in the group level as well as the subsidiary level because all data does not apply to all subsidiaries.

Because changes in the company impact data, changes in metadata should also be managed (Khatri & Brown 2010, 151). Referring to Smith (2007), Brous et al. (2016, 121) describe that the responsibility for metadata should be allocated to data stewards who have the necessary knowledge of the data under their stewardship. Therefore, it is important the responsibility for maintaining and updating metadata is assigned to a certain role to ensure that it stays up to date. As data stewards were assigned the responsibility for tactical scope of data management, it is seen as a suitable role for maintaining changes in metadata. Defining metadata directly addresses the need for standardized definitions for data terms and thus, facilitates effective communication about data issues. Metadata also connects financial data in understandable way to the daily operations, which is expected to increase common understanding about data in the company. Clear terms and common understanding provide the tools to communicate about data issues. Thus, it is expected to increase communication between the teams, which was identified as a challenge and a need for improvement among the interviewees. By specifying the uses of data, metadata also highlights the importance of the quality of financial data and hence, supports data strategy.

Data access

Data access is the fourth data governance domain and it refers to specifying access requirements for data. On one hand effective use of data requires that it is easily accessible. On the other hand, security of confidential data must be ensured by restricting who has access to certain data. (Khatri & Brown 2010, 149, 151.) Accessibility and security are also included in the data quality dimensions (Wang & Strong 1996, 14–15), which highlights the importance of data access decisions for data quality. Hence, data access policy is an integral part of data governance framework (Alhassan et al. 2019, 106). In the case company, data access was not identified as a challenge or a need for improvement. However, it is suggested that a data access policy is defined and documented in the company to create clear guidelines for data access requests. Additionally, data users' access rights to confidential data could be reviewed to confirm appropriate data security.

To prevent separate and possibly conflicting decisions of data access from being made, it is suggested that data access decisions are placed on one person. Because external regulations and norms influence a company's data access policy (Khatri & Brown 2010, 151) the norms regarding access of financial data should be considered in the case company. In the context of financial data this includes ensuring the confidentiality of salary data and complying to the guidelines that specify the segregation of duties. Segregation of duties is an element of internal controls that aims to prevent placing excessive control over a process on one person. Decisions of access rights require knowledge of the regulations of financial data as well as the data activities that data users conduct and thus, it is suggested that the decision-making authority for access rights is placed on the data owner. Because the data owner has the responsibility for the overview of financial data processes, the data owner has the necessary perspective for managing data access rights. Policy statements should be kept simple, understandable and up-to date to ensure that data stakeholders view them as beneficial rather than cumbersome (Alhassan et al. 2019, 106). Therefore, the suggested data access policy is that if an employee wants additional access to data, the approval should be asked from the data owner and after approval the access right should be requested from IT. This ensures that the access decisions regarding financial data are made in the finance function and IT does not have the primary responsibility for monitoring data access rights. The specified data access policy contributes to effective use of financial data when all data users have access to necessary data while simultaneously the confidentiality of data is ensured.

Data lifecycle

Data lifecycle is the fifth data governance domain and refers to defining the stages through which the data moves during its lifecycle. Managing data requires that a company is aware of what data it has, how critical that data is, where does the data come from and whether there are redundancies in the data. Defining data lifecycle includes mapping the data that the company owns and identifying the different life-cycle stages, such as production, retention and retirement of data. (Khatri & Brown 2010, 149–151). The case company had two needs regarding data lifecycle. Firstly, a need to define the existing financial data to clarify the amount of available data. Secondly, a need to define unclear data processes.

Weber et al. (2009, 12) mention identifying the data objects as a key task of data governance. In the case company, financial data had not been mapped, which hindered the understanding of the big picture of data. Therefore, it is suggested that the existing financial data is mapped and relevant data is identified and documented in the case company. When defining what data is relevant, the data strategy should be considered. Then, the sources from where the data comes from and the tasks where the data is used should be defined and added to the data map. Redundant data should be recognized from the data set and archived if the data is not found necessary for the company. Mapping and defining the data clarifies the amount of available data and ensures that irrelevant or redundant data does not interfere with effective data use. In addition, defining the available data contributes to understanding of the big picture of financial data for it to be effectively managed as a whole. Defining and documenting the uses of data contributes to data quality as data creators can better understand how the data is used. This is because data creators then have increased awareness of the requirements for data, because they have knowledge of the collection and utilization process of the data they use (Fletcher et al. 2005).

Clear data processes are a critical element of data governance framework and thus, they should be defined, implemented and monitored (Alhassan et al. 2019, 105). In the case company, data processes had not been defined, which was identified as a challenge for data quality. Especially the lifecycle of master data objects was found to be important to define to control the amount of available data and to ensure that the master data objects worked effectively together. Data processes should determine how data objects are created, maintained, utilized and deleted (Weber et al. 2009, 11–12). Because master data elements had a key role in financial data quality in the company, and new needs for master data were identified continuously, there should be a clear process for adding and deleting master data objects as well as defining who would make the decisions regarding that. Regarding the process for master data, it is suggested that decision making follows the structure that was defined regarding roles and responsibilities. The need for a new master data object may be identified by either data user, data steward or data owner. Then, data steward should make a proposal for the new data object to the data owner who approves the addition. Decisions should be made on the strategic level by data user who ensures that the addition fits to the data as a whole. After approval, the data steward is responsible for ensuring that the new data object is created, it is updated to metadata and it is communicated to all relevant data users. A clear process ensures that the big picture of data is considered in decisions and conflicting decisions over data are not made. Additionally, the policy clarifies the process flow to the data stakeholders. Regarding other lifecycle stages, such as making changes to or deleting master data objects, could be made by following the same process.

5 DISCUSSION AND CONCLUSIONS

The purpose of this research was to study how the quality of financial data can be improved by utilizing data governance. Financial data was defined as the data in the information system that directly affects a company's financial processes. The main research question that this paper aimed to answer was: *How can the data governance framework be utilized to address challenges in financial data quality*. The main research question was divided to two sub questions:

1) What kind of challenges regarding the quality of financial data exist?
2) How can the challenges be addressed with data governance framework?

The first research question was set to understand the factors that lead to defects in data quality. The necessary understanding of the topic for answering the research question was developed in the first section of the theoretical framework. First, data quality was discussed in the context of financial data and it was concluded that data quality should be evaluated in the context of the use (Wang & Strong 1996, 6). Therefore, financial data was defined as high quality when it fulfills the requirements of the task it is used for, such as financial accounting and management accounting. Then, barriers for high data quality were identified from the literature and they were classified to four themes which included management, roles and responsibilities, communication and technology related barriers. The empirical data for the research was collected from semi-structured interviews with the employees from financial accounting and management accounting teams in the case company. Semi-structured interviews provided the researcher the opportunity for open ended questions and redirecting the discussion during the interview and thus, enabled collecting in-depth information from the case company's challenges regarding financial data quality.

Management

Based on the empirical data, a key challenge related to management was the lack of leadership and overview for financial data. Financial data was not managed as a whole so it was not ensured that separate processes work effectively together and governance practices were under-developed. This supports the findings by O'Brien et al. (2013, 6) who observed that low priority of data governance created a challenge for data quality. Due to that, employees' awareness of data quality issues were low and data quality issues were not communicated in the organization. (O'Brien et al. 2013, 6.). Based on the interviews, lack of structure in data processes was found to cause unclarity among employees and lead to ad-hoc actions with data. This is in line with the study by Silvola et al. (2011, 155–157) who identified incoherent data management practices as a data quality barrier because they caused confusion among employees and made data maintenance a laborious task.

Insufficient controls identified in the literature (Haug & Arlbjørn 2010; Emeka-Nwokeji 2012; Xu 2015) were not explicitly mentioned as a data quality challenge in the case company. However, the lack of overall management practices implies that control practices were also under-developed. Tee et al. (2007) and Xu (2015) emphasize the role of management commitment for data quality. It was observed that employees conducting operational tasks often do not consider data quality in the bigger picture and therefore, the mandate to set data quality as a priority should come from the management. Lack of rewards was not perceived as a challenge to data quality in the case company, so this quality barrier identified by Haug & Arlbjørn (2010) was not identified in the empirical data. The challenges related to management in financial data quality were observed to be in line with the data quality barriers in the literature. This implies that management practices for data should be developed in organizations that want to achieve high financial data quality.

Roles and responsibilities

Based on the interviews, undefined roles and responsibilities were found to be a significant challenge to data quality. This supports the finding by Haug & Arlbjørn (2010) who found that the lack of delegation of responsibilities has the greatest negative impact on data quality. Because roles and responsibilities regarding financial data were not defined in the case company, it led to unclarity of the data users' responsibilities regarding data and ensuring data quality. When clear responsibility was not allocated, tasks affecting data quality were potentially left undone. This is in line with the finding by

O'Brien et al. (2013, 6) who observed that the employees' awareness of what is expected of them was negatively affected by under-developed structures and responsibilities regarding data. Silvola et al. (2011, 155–157) found undefined data ownership as one of the most common challenges related to data quality. In the case company, challenges from undefined data ownership were observed as unclear decision-making authority as the interviewees described uncertainty in decision-making regarding data. Undefined decision-making right created a risk that all data stakeholders were not taken into account. In addition, without overview of financial data, individually made decisions created a risk for data consistency. Haug & Arlbjørn (2010, 301) note that while defining responsibilities is a relatively easy and affordable way to address these challenges, it can have a substantial positive effect on data quality as it creates clear accountability over data. (Haug & Arlbjørn 2010, 301.)

Communication

Communication was found to have a material effect on financial data quality as all of the interviewees described challenges due to insufficient communication. The lack of communication between different teams using financial data hindered common understanding of how data is used in the organization and thus, impeded data creators' awareness of the requirements for data. This is in line with the study of Fletcher et al. (2005) who found that knowledge of the collection and utilization process of data enhances data creators' awareness of the requirements for data and thus, supports achieving high data quality. Therefore, companies should emphasize communication between data stakeholders to facilitate common understanding of the data processes in the organization. In addition to increasing awareness of data requirements, this would contribute to understanding of the different processes in which financial data is used in the company, which is needed to manage financial data as a whole.

Tee et al. (2007, 351) found awareness of data quality to be an important factor for data quality. In the context of financial data, communicating data quality requirements to the data creators was found especially important because employees without financial background are not aware of the requirements that external regulations set for financial data. Therefore, finance personnel should be able to translate the data requirements organization-wide in an understandable way to ensure compliance. In this vein, training and written instructions may be beneficial for communicating the needs for financial data to the data stakeholders. Also, because financial data is stored and used in information systems, it highlights the importance of IT competencies. Knauer et al. (2020, 102) suggest that internal IT competencies should be developed in all organization's functions that operate with data, not only in the IT function. Communication was also found to be negatively affected due to different understanding of the terms used in communication. The lack of common understanding was found to lead to misunderstandings and thus, negatively affect data quality. This finding is in line the study of Silvola et al. (2011) who found that unclear data definitions caused problems for communication and thus, lead to decreased data quality. This implies that unambiguous terms are necessary for communicating effectively about data. Effective communication also further facilitates common understanding of data in an organization.

Technology

Based on the empirical data, manual tasks in the information system were seen as a challenge to financial data quality. This is because they increase the possibility of data defects due to human error especially in complex tasks. This is consistent with the study by Xu (2015) who found that nature of accounting information system has a key role for financial data quality. This includes that the system is easy to use, a high degree of data validation is automated and that the system is up to date (Xu 2015, 9). Hence, the possibility of human error could be limited with automation. However, all tasks with financial data cannot be fully automated because they often require some degree of human assessment. Also, Haug & Arlbjørn (2010, 300) mention the impact of user-friendliness of the software on data quality. In the case company, it was observed that the information system was unable to conform to all requirements that had been set for the data. Thus, defects in financial data may also be due to the characteristics of the information system.

Silvola et al. (2011, 160) mention integrations between applications as a data quality barrier. They observed that challenges to data quality arise when data is transferred from one system to other. This challenge was also observed by one interviewee who mentioned that due to integrated systems, all necessary fields of customer data did not flow to the information system where financial data was eventually stored. In conclusion, challenges due to technology were found to be a minority in the case company. This was partially due to the fact the research focused on organizational data quality challenges. However, technical factors also have a role in achieving high data quality as information system provides the infrastructure where data activities are conducted (Cheong & Chang 2007, 1002).

Internal conditions

Challenges that were categorized as internal conditions included factors in the company's internal situation that created challenges for achieving high data quality. These challenges included scarce resources, fast changes in the company and frequently changing data requirements which were mostly due to the company's fast growth during the past years. These challenges were not currently the key challenges in the case company but they had previously affected data quality. The mentioned challenges are highly context specific but they imply that the factors in the company's internal conditions also have an effect on data quality. Therefore, these internal factors should also be evaluated among the four other themes for companies to understand the effect of internal factors on data quality. A theme of challenges related to internal conditions was not identified in the theoretical framework of the study. However, Xu (2015, 7) mentions organizational structure and culture as factors impacting data quality based on previous literature, which refer to internal factors. Based on her empirical data, she did not find these factors as the most important for accounting information system data quality (Xu 2015, 1).

Internal conditions were not found to be the most important factors for financial data quality. However, they are relevant factors to consider because without adequate attention, these factors that will negatively affect data quality even if structures for responsibilities and managing data are established in a company. Thus, these factors cannot be directly prevented from incurring but they highlight the internal conditions that need to be considered when aiming to improve data quality. Firstly, adequate resources should be ensured for maintaining data quality and the importance of data quality should be emphasized. Secondly, data management practices need to be evaluated regularly and updated if changes have been made in the company. Thirdly, management practices must enable flexible changes to data while ensuring that data quality remains on an appropriate level.

In conclusion, the challenges for data quality in financial data that were identified from the interviews are in line with the data quality barriers identified in the literature. Challenges were identified from all four themes used in the theoretical framework, which included management, roles and responsibilities, communication and technology. In addition, challenges related to internal conditions was discussed as a theme that was not identified in the literature in the theoretical framework. Based on the empirical data, these factors can decrease data quality and thus, they should be addressed in order to achieve high quality in financial data.

Data governance framework

The second research question was set to examine, how can data governance framework be utilized to address the data quality challenges in financial data and thus, to contribute to high data quality. Data governance was chosen as the tool because it is seen as a promising way to address organizational data quality issues (Benfeldt Nielsen 2017, 120). The theoretical framework consisted of data governance literature clarifying the concept and introducing two data governance models. Following a design-based research approach, an answer for the second research question was formed by developing a data governance framework for the case company. The framework was developed based on the contingency model by Weber et al. (2009) and data governance model by Khatri and Brown (2010). The aim was to address the data quality challenges that were identified in the interview data and to fulfill the improvement needs that were mentioned by the interviewees.

First, the data governance roles were defined to address the challenges related to roles and responsibilities and create a foundation for data governance activities. The suggested framework consisted of three levels which were strategic, tactical and operational level. In the case company the challenge regarding roles and responsibilities was that they had not been defined which created unclarity of decision-making. In addition, defining roles and responsibilities, such as data ownership was mentioned as a key need for improvement by the interviewees. Therefore, the structure of roles and responsibilities was designed to address this challenge. Firstly, defining a role for each data stakeholder clarifies their influence on financial data quality. Secondly, allocating responsibilities for each role aims to minimize data defects due to collective responsibility and undone tasks.

Thirdly, because uncoordinated decision-making was found to create a risk for data quality, decision-making authority was included in the suggested data governance framework to address this risk.

Data governance activities in the suggested framework followed the data governance model by Khatri and Brown (2010), which included data principles, data quality, metadata, data access and data lifecycle. An activity was suggested for each of these domains to ensure that all critical aspects of data governance are covered in the framework. In the domain of data access, challenges or needs for improvement were not identified in the empirical data. However, it was suggested that data users' access rights are reviewed and a data access policy for financial data is established to ensure the confidentiality of data.

In the suggested framework, data principles and data lifecycle domains were especially focused on addressing the management related data quality challenges. In the domain of data principles, it was suggested that the company forms a data strategy for defining the strategic goals that it aims to pursue with financial data. These goals are further aligned with the company's strategic goals to highlight the importance of data quality (Alhassan et al. 2019, 107). Because the lack of leadership of financial data as a whole was identified as a data quality challenge, data strategy is needed to clarify the direction for managing financial data quality. Connecting data to the company's strategy also clarifies the strategic importance of data as a company asset for the whole organization (Khatri & Brown 2010, 150). Data strategy also highlights the need for managing financial data and thus, may contribute to the adoption of data governance. In addition, data strategy would work as a tool for communicating data quality issues in the organization and thus, also mitigate the challenge of communication.

Data lifecycle referred to understanding how data moves through different stages during its lifecycle (Khatri & Brown 2010, 151). In the case company, financial data had not been mapped which affected the understanding of the big picture of financial data. The lack of structure in data processes was due to the fact that the processes had not been defined. In addition, defining data processes and mapping all financial data were mentioned as needs for improvement by the interviewees. Therefore, to address these challenges and to gain understanding of the data lifecycle, it was suggested that all financial data is mapped and relevant data is identified (Khatri & Brown 2010, 151). In addition, it was suggested that data processes are defined and a proposal for the process of making changes to master data was presented. This would create clarity for conducting the data processes and reduce the need for ad-hoc actions which in turn would positively impact data quality. Understanding of the data processes would also contribute to data creators' understanding of the requirements for data (Fletcher et al. 2005). In addition, understanding the relationships of data processes would be necessary for managing financial data as a whole.

Data quality and metadata domains focused particularly for addressing challenges in communication. Regarding data quality, following Khatri and Brown (2010, 150) it was suggested that data quality requirements for financial data were clearly defined and documented. In terms of financial data, these definitions should include the requirements that external regulations set for the data. Most importantly the defined data quality requirements should be communicated to the data stakeholders (Alhassan et al. 2019, 107). This is essential to ensure that the data creators have adequate awareness of the data quality requirements for the data they produce. This, data quality domain addresses the lack of awareness of data quality that was identified as a challenge in the case company. In addition, this contributes to the management of financial data as the requirements for data quality are clear.

Misunderstandings in communication were also identified as a challenge to data quality and thus, a need to create clear definitions for financial data was recognized. Following Khatri and Brown (2010, 150), establishing metadata was suggested to mitigate this challenge by creating standardized definitions for data for the whole company. Especially regarding domain-specific financial data, metadata that includes unambiguous definitions for data related terms, can be used to communicate finance needs to the company in an understandable way. Metadata contributes to common understanding and thus, facilitates effective communication among data stakeholders.

In conclusion, it seems that data governance is suitable tool improving data quality in the case company because it mitigates the organizational data quality challenges that were identified in the interviews. However, the effectiveness of the data governance framework highly depends on how well it is implemented and adopted in the company. This is because addressing these data quality challenges requires a cultural change to the way people think of financial data and use it.

It was also observed that the characteristics of financial data should be considered when designing a data governance framework. Following the decision domains by Khatri and Brown (2010), firstly, the nature of financial data influences the strategic goals that are set to be pursued with the data. Secondly, the requirements that external regulations set for financial data should be taken into account when definitions for data quality requirements are formed. Thirdly, as the terms used for financial terms are often context specific, they should be defined in an understandable way organization wide to facilitate common understanding. Fourthly, data access practices should consider the guidelines that are set for confidential financial information and segregation of duties in companies. Fifthly, as financial data may be widely used in companies in different functions, the complexity should be mitigated by defining the financial data and data processes to create awareness of their interrelatedness.

Conclusions and future research possibilities

The findings of this research have academic and practical contributions. In terms of academic contributions, this research extends the understanding of data quality challenges and their effects on data quality in the context of financial data. A classification of five themes of data quality challenges was created, which provides a framework for further examination of data quality barriers. In addition, this research contributes to the limited literature of utilizing data governance in the context of financial data by giving an example of designing data governance framework for financial data and specifying the characteristics of financial data that should be considered when designing a data governance framework.

As a practical contribution, the findings of this research can help practitioners to better identify and understand data quality challenges in financial data. The classification of data quality challenges may be useful for practitioners for examining data quality barriers in organizations. In addition, the study clarifies the use of data governance as a possible method to mitigate data quality challenges in financial data. By presenting a case example of designing a data governance framework for financial data, this research aims to provide an understandable example of the concept. Practitioners can use this as a case example in designing their own data governance activities. Most importantly, this research aims to motivate accounting professionals to pay attention to and take responsibility for the quality of financial data in organizations.

The findings of this research are subject to certain limitations. The nature of case study is studying one company in order to gain in-depth understanding of the phenomenon (Hirsjärvi et al. 2009, 130). Therefore, the characteristics of the case company influence the findings and the findings are tied to this particular context. Additionally, the empirical data was collected in interviews meaning that the findings are interpretations of the interviewees' subjective experiences. Therefore, the findings cannot be generalized to other companies. In addition, typically the end state of a design-based study is not a suggestion to the case company but rather the suggested solution should be implemented and its effectiveness should be evaluated (Tamminen 1993, 158). Due to the time scope of this research, evaluating the effects of implementing data governance framework was not possible. Therefore, this research does not fulfill all of the requirements of a design-based study.

The results of this study suggest several possibilities for future research. Examining the effects of the implementation of the data governance framework was not possible in the scope of this research. Therefore, a longitudinal study would enable collecting evidence of the effects of the data governance framework on financial data quality. In this vein, the factors that affect the implementation of data governance could also be examined. Additionally, more research could be done on alternative methods, such as quantitative methods to contribute to the understanding of the phenomenon. A quantitative study of data governance in the context of financial data could be conducted as a survey to study how widely data governance practices are used in companies for governing financial data. Alternatively, a research focusing on the data quality barriers in financial data as individual factors could provide in-depth understanding for directing data governance activities specifically for financial data. More research of data governance in the context of financial data is needed because of the critical role that high-quality financial data has on companies' success.

REFERENCES

Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management,* 49, 424–438.

Abueed, R., & Aga, M. (2019). Sustainable Knowledge Creation and Corporate Outcomes: Does Corporate Data Governance Matter? *Sustainability*, 11(20), 1–15.

Ackoff, R. L. (1999). *Ackoff's Best*. New York: John Wiley & Sons, 170–172.

Alhassan, I., Sammon, D., & Daly, M. (2019). Critical Success Factors for Data Governance: A Theory Building Approach. *Information Systems Management*, 36(2), 98–110.

Allen, M. & Cervo, D. (2015). *Multi-domain master data management: Advanced MDM and Data Governance in Practice*. Morgan Kaufmann. USA.

Alpar, P., & Winkelsträter, S. (2014). Assessment of data quality in accounting data with association rules. *Expert Systems with Applications*, 41(5), 2259–2268.

Atkinson, K., & McGaughey, R. (2006). Accounting for data: a shortcoming in accounting for intangible assets. *Academy of Accounting and Financial Studies Journal*, 10(2), 85–95.

Aula, R. (2020). Mitä data governance tarkoittaa ja entä, jos sitä ei olisi olemassa? Accessed 15.10.2020. https://www.tietoevry.com/fi/uutishuone/kaikki-uutiset-ja-tiedotteet/blogi/2020/mita-data-governance-tarkoittaa-ja-enta-jos-sita-ei-olisi-olemassa/

Bai, X., Nunez, M., & Kalagnanam, J. R. (2012). Managing Data Quality Risk in Accounting Information Systems. *Information Systems Research*, 23(2), 453–473.

Bhimani, A., Horngren, C., Datar, S., & Rajan, M. (2015). *Management and cost accounting* (6.ed.). Pearson.

Brous, P., Janssen, M., & Vilminko-Heikkinen, R. (2016). Coordinating Decision Making in Data Management Activities: A Systematic Review of Data Governance Principles. *Electronic Government*, 9820, 115–125.

Bryman, A., & Bell, E. (2015). *Business research methods* (4.ed.). Oxford University Press.

Cao, L., & Zhu, H. (2013). Normal accidents: Data quality problems in ERP-enabled manufacturing. *Journal of Data and Information Quality* (JDIQ), 4(3), 1–26.

Cheong, L. K., & Chang, V. (2007). The need for data governance: A case study. *Proceedings of the 18th Australasian Conference on Information Systems ACIS 2007*, 999–1008.

Du, J., & Zhou, L. (2012). Improving financial data quality using ontologies. *Decision Support Systems,* 54(1), 76–86.

Dreibelbis, A. (2008). *Enterprise master data management: an SOA approach to managing core information*. IBM Press/Pearson plc.

Emeka-Nwokeji (2012). Repositioning Accounting Information System Through Effective Data Quality Management: A Framework For Reducing Costs And Improving Performance. *International Journal of Scientific & Technology Research*, 1, 86–94.

Even, A., & Shankaranarayanan, G. (2007). Utility-driven assessment of data quality. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, 38(2), 75–93.

Fletcher, D., Robbert, M., Mohamad, K., & Middleton, P. (2005). A systems approach to monitoring financial data quality assessment and improvement. *Proceedings of the 2005 International Conference on Information Quality, ICIQ 2005.*

Framery (2020). Framery webpage. Accessed 3.11.2020. https://www.frameryacoustics.com/en/company/

Framery (2021). Framery sustainability report 2020. Accessed 21.4.2021 https://www.frameryacoustics.com/en/company/sustainability/

Glowalla, P., & Sunyaev, A. (2014). ERP system fit – an explorative task and data quality perspective. *Journal of Enterprise Information Management*, 27(5), 668–686.

Granlund, M., & Malmi, T. (2002). Moderate impact of ERPS on management accounting: a lag or permanent outcome? *Management Accounting Research*, 13(3), 299–321.

Haug, A., Arlbjørn, S.J., & Pedersen, A. (2009). A classification model of ERP system data quality. *Industrial Management + Data Systems*, 109(8), 1053–1068.

Haug, A., & Arlbjørn, S.J. (2010). Barriers to master data quality. *Journal of Enterprise Information Management*, 24(3), 288–303.

Haug, A., Zachariassen, F., & van Liempd, D. (2011). The costs of poor data quality. *Journal of Industrial Engineering and Management,* 4(2), 168–193.

Hirsjärvi, S., Remes, P., & Sajavaara, P. (2009). *Tutki ja kirjoita* (15.ed.). Tammi.

Kamioka, T., Luo, X., & Tapanainen, T. (2016). An empirical investigation of data governance: The role of accountabilities. *PACIS 2016 Proceedings*. 1–12.

Kasanen, E., Lukka, K., & Siitonen, A. (1991). Konstruktiivinen tutkimus liiketaloustieteessä, *Liiketaloudellinen aikakauskirja*, vol. 3, 301–325.

Khatri, V., & Brown, C. (2010). Designing data governance. *Communications of the ACM*, 53(1), 148–152.

Kihn, L. A., (2015). Laatu kolmesta tarkastelunäkökulmasta: tekninen, kaupallinen ja palveluun liittyvä laatu. Teoksessa af Ursin, K., Pekkola, E. & Stenvall, J. (toim.) *Felix byrokratia? Julkinen hallinto kaiken huomioimisen taitona*. Tampere: Tampere University Press, 283–301.

Kihn, L. A. & Ihantola E-M. (2008). Tutkimuksen laadun arvioinnista. Teoksessa: Hyvönen, T., Laine, M. & Mäkelä, H. *Laskenta-ajattelun tutkija ja kehittäjä Professori Salme Näsi 60 vuotta*. Tampereen yliopisto, Tampere, 81–95.

Kihn, L. A., & Näsi, S. (2010). Research strategic analysis of the Finnish doctoral dissertations in management accounting from 1990 to 2009. *Liiketaloudellinen Aikakauskirja*, 59(1), 6–7, 41–85.

Kirjanpitolaki 30.12.1997/1336

Knauer, T., Nikiforow, N., & Wagener, S. (2020). Determinants of information system quality and data quality in management accounting. *Journal of Management Control*, 31(1-2), 97–121.

Kooper, M., Maes, R., & Lindgreen, E. (2011). On the governance of information: Introducing a new concept of governance to support the management of information. *International Journal of Information Management*, 31(3), 195–200.

Koskinen, I., Alasuutari, P. & Peltonen, T. (2005). *Laadulliset menetelmät kauppatieteissä*. Vastapaino, Jyväskylä.

KPMG (2020). Guardians of trust. Accessed 1.11.2020. https://home.kpmg/content/dam/kpmg/xx/pdf/2018/02/guardians-of-trust.pdf

Ladley, J. (2012). *Data governance how to design, deploy, and sustain an effective data governance program*. Morgan Kaufmann.

Laki oma-aloitteisten verojen verotusmenettelystä 9.9.2016/768

Liu, Q., Feng, G., Zhao, X., & Wang, W. (2020). Minimizing the data quality problem of information systems: A process-based method. *Decision Support Systems,* 137, 1–12.

Malinić, S., & Todorović, M. (2012). How Does Management Accounting Change under the Influence of ERP? *Economic Research-Ekonomska Istraživanja*, 25(3), 722–751.

Mosley, M. (2008). *The DAMA dictionary of data management*, (1.ed.). Technics Publications.

Benfeldt Nielsen, O. (2017). A Comprehensive Review of Data Governance Literature. *Selected Papers of the IRIS*, 8(3), 120–133.

Neilimo, K., & Näsi, J. (1980). *Nomoteettinen tutkimusote ja suomalainen yrityksen taloustiede: tutkimus positivismin soveltamisesta*. Tampereen yliopisto.

O'Brien, T. (2015). "Accounting" for Data Quality in Enterprise Systems. *Procedia Computer Science*, 64, 442–449.

O'Brien, T., Sukumar, A., and Helfert, M. (2013). The Value of Good Data - A Quality Perspective. *The International Conference of Enterprise Information Systems.* Angers, France

Otto, B. (2011a). Data Governance. *Business & Information Systems Engineering*, 3(4), 241–244.

Otto, B. (2011b). A morphology of the organisation of data governance. *19th European Conference on Information Systems, ECIS 2011.*

Pierce, E., Dismute, WS., & Yonke, CL. (2008). The state of information and data governance – understanding how organizations govern their information and data assets. *Information and Data Governance Report. IAIDQ and UALR-IQ*

Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research methods for business students* (5.ed.). Prentice Hall.

Schreider, T. (2020). *Cybersecurity Law, Standards and Regulations,* (2.ed). Rothstein Associates, Incorporated.

Silvola, R., Jääskeläinen, O., Kropsu-Vehkaperä, H., & Haapasalo, H. (2011). Managing one master data - challenges and preconditions. *Industrial Management + Data Systems,* 111(1), 146–162.

Strong, D., Lee, Y., & Wang, R. (1997). Data Quality in Context. *Communications of the ACM,* 40(4), 103–110.

Tallon, P., Ramirez, R., & Short, J. (2013). The Information Artifact in IT Governance: Toward a Theory of Information Governance. *Journal of Management Information Systems*, 30(3), 141–178.

Tamminen, R. (1993). *Tiedettä tekemään*. Atena.

Tayi, G., & Ballou, D. (1998). Examining data quality. *Communications of the ACM*, 41(2), 54–57.

Tee, S., Bowen, P., Doyle, P., & Rohde, F. (2007). Factors influencing organizations to improve data quality in their information systems. *Accounting and Finance (Parkville),* 47(2), 335–355.

Todd, G. (2008). Data Governance: The Enabler of High Performance. *DM Review*, 18(5), 30–33.

Tuomi, J., & Sarajärvi, A. (2018). *Laadullinen tutkimus ja sisällönanalyysi* (Uudistettu laitos.). Tammi.

Vayghan, J. A., Garfinkle, S. M., Walenta, C., Healy, D.C., & Valentin, Z. (2007). The internal information transformation of IBM. *IBM Systems Journal*, 46(4), 669-684.

Wang, R., & Strong, D. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12(4), 5–33.

Warren, J., Moffitt, K., & Byrnes, P. (2015). How big data will change accounting. *Accounting Horizons*, 29(2), 397–407.

Watson, H., Fuller, C., & Ariyachandra, T. (2004). Data Warehouse Governance: Best Practices at Blue Cross and Blue Shield of North Carolina. *Decision Support Systems,* 38(3), 435–450.

Weber, K., Otto, B., & Österle, H. (2009). One Size Does Not Fit All---A Contingency Approach to Data Governance. *Journal of Data and Information Quality (JDIQ)*, 1(1), 1–27.

Wende, K. (2007). A Model for Data Governance – Organizing Accountabilities for Data Quality Management. *Proceedings of the 18th Australasian Conference on Information Systems ACIS 2007*, 417–425.

Xu, H. (2015). What are the most important factors for accounting information quality and their impact on AIS data quality outcomes? *ACM Journal of Data and Information Quality,* 5(4), Article 14, 1–22.

APPENDIX 1: INTERVIEW QUESTIONS

Introduction

1. What is your role in the company?
2. How do you understand the term financial data?
3. What is the role of financial data in your tasks?
4. How do you understand the term data quality?
5. How is data quality visible in your work?

Financial data quality

6. How important are financial data and data quality to Framery in your opinion?
7. How do you perceive the level of quality of the data you use?
8. What kind of challenges have you faced regarding data quality in your work?
9. Can you give examples of situations where you faced problems with data quality?
10. What have been the causes for the data quality challenges in your opinion?
11. How do you think the problems with data quality affect your work?
12. Have you noticed data quality challenges outside of your own work?
13. Have you recognized possible risks due to data quality challenges?

Data governance

14. Do you find that data usage and management are organized at Framery?
15. Do you find that clear processes around data exist?
16. Are responsibilities between different data users defined and clear?
17. Do you think that there is enough communication between different data users?
18. Do you think that people are aware of how their actions affect data quality?

Improvements

19. What kind of needs do you recognize regarding data quality and data management?
20. How do you think you would benefit from better managed financial data?
21. How do you think it would affect your tasks?
22. How do you think the company would benefit?