WIRTSCHAFTS
UNIVERSITÄT
WIEN VIENNA
UNIVERSITY OF
ECONOMICS
AND BUSINESS

Bachelor Thesis

Challenges and opportunities of Data Governance in private and public organizations

Christian Bruck

Date of Birth: 04.03.1996
Student ID: h01550727

Subject Area: Information Business

Studienkennzahl: J033561

Supervisors: Univ.Prof. Dr. Axel Polleres and
Univ.Prof. Dr. Edward W. Bernroider

Date of Submission: 29 October 2017

Department of Information Systems and Operations, Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, Austria

DEPARTMENT FÜR INFORMATIONS-
VERARBEITUNG UND PROZESS-
MANAGEMENT DEPARTMENT
OF INFORMATION SYSTEMS AND
OPERATIONS

We don't have better algorithms, we just have more data.
Peter Norvig, Director of Research at Google

Contents

List of Figures

List of Tables

Abstract

In times of digitalization and globalization, organizations are facing numerous data-related issues. Since the amount of available data increases steadily, both private and public organizations are concerned with managing their valuable data assets properly according to their specific requirements. The concept of Data Governance encourages the achievement of organizational objectives by developing and implementing a suitable multidimensional data strategy on methodological, organizational and technical levels. Aligned with a suitable implementation, organizations are facing a number of challenges while also achieving considerable benefits. It is the goal of this thesis to identify driving factors for Data Governance as well as to take a deeper look at different approaches on how to cope with this topic. The majority of the identified challenges and opportunities are applicable to each organizational type, and therefore seem to generally support decision-makers in establishing their Data Governance strategy. It can be concluded that a complete Data Governance program affects organizations of each size and can deliver benefits in either case. Large companies are already concerned with these concepts and therefore relatively advanced. On the next level, many SMEs have also recognized the promising potential of this data-driven perspective and align their business strategy to its main ideas, but more implicit and less formal than in the first category. In case of the public organizations, a Data Governance approach is also appreciated but regarded more critically as it is complicated to provide an appropriate cost-benefit weighing. Overall, benefits are likely to derive for each organization considering their individual circumstances, whereas various regulatory challenges are needed to be faced in order to satisfy the different parties involved and influenced by such a Data Governance program. The insights resulted from a preliminary literature review, subsequent semi-structured interviews and a final online survey, and relate to organizations from Austria and Germany. Overall, the focus has been put on combining business and IT-related perspectives and processes in order to reach a broader view on the available resources and thus to enable these organizations to recognize and achieve their maximum potential for further business opportunities in the near future.

Acknowledgements

First of all, I would like to thank my supervisors Univ.Prof. Dr. Axel Polleres and Univ.Prof. Dr. Edward W. Bernroider for their continuous support and meetings during the whole process of gaining the relevant material and writing down the results. I am thankful for having the chance of developing my bachelor thesis in this comprehensive and future-oriented domain which affects organizations of each type sooner or later. Furthermore, I would like to thank my whole family for their encouragement in difficult situations and understanding that generating this thesis has required numerous resources. My final thanks goes to all interview partners and participants of the online survey who have played a major part in generating the material necessary for the ultimate outcome of this thesis.

1 Introduction

Data Governance is a comprehensive concept. Due to the increasing complexity of the business world, data as a value-creating asset within different types of organizations need to be governed accordingly. In this context, Data Governance as an emerging trend in enterprise information management has been identified as being a potential tool for supporting this process.

By now, numerous models and frameworks exist, leaving companies and public institutions of various sizes with many ways to choose their own direction in this development. These tools have mainly evolved in the American environment and are therefore not directly applicable for European countries. Consequently, these organizations are facing difficulties in implementing a convenient Data Governance strategy.

The motivation to conduct further research in this area is related to the fact that Data Governance represents an increasingly essential aspect in organizations but has not been widely surveyed yet, especially not for Austrian companies and public organizations. It also seems to be the "right time" to conduct a deepening analysis in this field since numerous companies are at the moment trying to adapt their strategies and business model towards a more data-oriented approach in order to achieve an overall competitive advantage. Therefore, the current situation in Austria and Germany has been examined by identifying existing challenges and opportunities of Data Governance. The information gained should support the implementation of new data-driven business strategies for organizations within the near future. A particular focus has been put on the various and current forms of implementation of Data Governance inside the companies and institutions in order to achieve a better understanding of its key concepts.

As outlined below, there exists a considerable number of definitions for the term Data Governance. One reason is that each company has to implement its own specific Data Governance strategy over time. As a result, these realizations can differentiate from each other, even within a single subsidiary of a company as they can be of different size, company culture or business strategy. Thus, a various understanding of this concept evolves and the claims for standardized frameworks are increasing. Consequently, it is one goal of this thesis to outline a common definition of Data Governance in private and public organizations in German-speaking countries, especially in Austria and Germany by taking a closer look at their current realization.

The following section considers the formulated research question and the methodological research approach. It also consists of definitions for the research entities and argues limitations of the research method. The literature review that follows broadly introduces the reader to the topic and offers a

suitable definition of Data Governance for this thesis. The main part covers the findings derived from the interviews. In the final part, some major results of the online questionnaire are considered, followed by resulting conclusions and future work. The appendix is split into the main tables with the results from the interviews and the online survey. Moreover, it also includes additional graphics concerning the results of interviews and a blank version of the complete interview and online questionnaire.

1.1 Research question

The aim of this thesis is to analyze the current situation of Data Governance in large companies, public organizations and small and medium-sized enterprises (SMEs) in German-speaking countries with regard to different industries, by giving a general overview of potential implementations and to identify various challenges and opportunities which are part of this process. Given this context, the defined research question is:

"What are challenges and opportunities deriving from the implementation of a Data Governance strategy within private and public organizations in German-speaking countries?"

In order to find a suitable answer to this research question, a number of sub-questions have been formulated:

- How do organizations deal with their data?
- What are common organizational roles of a Data Governance program?
- Have concepts such as "Big Data", "Data Spaces", "Data Lakes" or "Enterprise Linked Data" already arrived within the organizations?
- What role do data in corporate strategy play?
- What new challenges, opportunities and issues have emerged from the flood of data in a company and how is it managed internally by these organizations?
- Are the roles of a Chief Data Officer (CDO) and Data Scientists already implemented within organizations?

The hypothesis is that there are challenges and opportunities which differ from the specific type of the organization. A particular focus is being put on the question how far companies and institutions are aware of these different

concepts and how they are implemented in relation to their Data Governance program. The analytical part of the research method focuses on these aspects as well as on resulting challenges and opportunities. Consequently, the overall focus is mainly put on organizational rather than on technical issues. Moreover, the importance of different dimensions of a specific Data Governance framework are considered. Last but not least, the thesis will observe companies and institutions of different industries and will therefore not focus on a particular one.

1.2 Research methodology

In order to find an answer to the research question, a three-tiered approach has been chosen. In a nutshell, the conducted research can be described as qualitative, descriptive and exploratory. It included a precise literature review in the first place which consisted of a topic review and a case study method. In parallel to this early research state, the interview questions for the subsequent interview survey had been formulated and adapted to the findings from the literature review. The following qualitative part focused on the interview survey, and these results were then used for the conduction of a final online survey. The focus has been put on three types of organizations, namely large companies, small and medium sized enterprises (SMEs) and public institutions. Their definitions are outlined in the next section. In total, nine cases have been observed during the research process in the interviews. It should be mentioned that all interview partners have been aware of IT related concepts and tools to a relatively high extent.

Table 1: The nine cases in chronological order by the date of the interview.

Date	Type of Organization	Case No.	Purpose of Case
Sept 1, 2017	Public organization	Case 1	Description
Sept 6, 2017	Public organization	Case 2	Theory-testing
Sept 6, 2017	Large company	Case 3	Description
Sept 8, 2017	Public organization	Case 4	Theory-testing
Sept 11, 2017	Large company	Case 5	Description
Sept 11, 2017	SME	Case 6	Description
Sept 14, 2017	Large company	Case 7	Theory-testing
Sept 14, 2017	SME	Case 8	Theory-testing
Sept 20, 2017	SME	Case 9	Theory-testing

With the use of data triangulation as a research methodology, the situation within the first of each of the three categories has been identified at the

beginning. This information has been then used within the second case to find similarities between them. Although some have been identified within the selected nine cases, they are rather conceptually distinct. As far as the public organizations and the SMEs are concerned, the first case has been used to describe the situation, whereas both subsequent cases were used for testing this theory. Only in case of the large companies the first two interviews both had a descriptive purpose with testing this theory in the third case, which was due to the amount of new information gained from the second case.

The analysis of the fieldwork included a pattern-matching approach within and across the cases, separated by their organizational type in a first step [36]. Thereafter, the identified challenges and opportunities have been summarized by their appearance within the interviews and the existing literature, before being categorized by specific criteria, which have been identified during the analytical process.

As for the creation of the interview questionnaire, the guidelines for a qualitative research in Information Systems research have been used [48]. The opening, introduction, key questions and close of the script have been prepared in advance.

The considered Data Governance model - the "Data Management Body of Knowledge" (DMBOK) - has been chosen due to the fact that it provides an easy to understand but detailed enough overview of important areas involved in a Data Governance program and because of a lack of a suitable framework for German-speaking countries. This model consist of eleven different dimensions of data management with Data Governance as the central point, which are divided into further sub-sections. As a first step, these ten surrounding dimensions have been analyzed for specific concepts, which turned into research objects for the interview questions. It should be remarked that the terms of "privacy and data protection" were added to the dimension data security since this is an essential part but not explicitly mentioned in this model.

1.2.1 Organizational types

Within the existing literature, organizations are divided into three common organizational types, depending on aspects such as the total number of employees. These three categories "SMEs", "public organizations" and "large companies" are defined as follows:

Medium-sized enterprises are characterized as enterprises which employ not more than 249 persons, have an annual turnover of no more than 50 million euro, and/or an annual balance sheet with not more than 43 million euro in total. Small-sized enterprises employ up to 49 persons, with an annual

turnover of not more than 10 million euro, and/or an annual balance sheet which sum does not exceed 10 million euro. Micro-sized enterprises employ less than 10 persons and have an annual turnover and/or an annual balance sheet of 2 million euro at the most. These figures have been valid since January 1, 2005 and represent maximum limits [76].

On the other hand, public enterprises, or state-owned enterprises, are organizations which have publicly-owned undertakings or privately-owned undertakings with at least a majority of ownership by the state [17]. For this thesis, we add public institutions (i.e. public enterprises without a specific legal form) to this type of research object and summarize these two organizational types under the term "public organization". These include for example institutions of the federal administration, authorities, promotion agencies, educational and research institutions, interest representatives and social insurance funds.

Large companies are consequently all other companies that do not fall into these categories and are no public organizations.

1.2.2 Literature review

As for the literature review, a topic review was chosen, which has focused on search keywords for a specific topic within a relatively broad field. This type of literature review is conceptual-content-related rather than focusing on specific studies or theories. The selection criteria for the search filters have been defined in advance and derived mainly from subsequent aspects on how to answer the main research question, i.e. by splitting the term Data Governance into some of its major elements. In fact, the search terms were rather general since detailed scientific surveys on rather specific issues have not been conducted so far. The complete table of this process with the used keywords and search strings can be found in the appendix. Articles have been collected from different online databases with the help of appropriate search engines, which focused on Google Scholar from which other ones such as Springer Link, Science Direct and ACM Digital Library were addressed.

The selected articles were sorted by their appropriateness by title, abstract and conclusion. In some of the relevant papers, the references were examined as well. Although "snowball-search" was practiced it should be said that due to the limited number of topic-related papers available, most of the references link to the same main papers. The chosen papers include scholar journals, presented papers and edited books. This search strategy also includes white papers from different organizations. We suppose that they are useful for answering the research question since they were conducted and released by organizations which tried to survey their immediate surrounding

and thus providing a useful view on the topic which has not been covered in some research papers by now. The information gained from this non-scientific literature will be separated explicitly from the scientific facts of the research papers.

In general, it could be stated that there is not much literature available, especially not for specific economic cases or industries. Additionally, most of the few papers available focus on the situation in the US and do not consider any European country. However, some of the papers have not been selected in case the focus was not put in enough detail on Data Governance itself (e.g. Data Governance in healthcare). Moreover, some of the white papers were not used for this work if their focus was rather on the qualities of specific tools or products for the use within a Data Governance program instead of considering the theoretical concepts. In the end, we have chosen 59 publications and 13 white papers for the examination of the state-of-the-art, with a particular focus on the identified challenges and opportunities of Data Governance.

1.2.3 Interviews

A total number of nine interviews with different Austrian organizations were carefully planned and conducted for the gain of the empirical material for this thesis. Each of the three selected types of organizations has been represented by three different interview partners who belong to a different industry. It is important to note that no organization which is directly responsible for the development of software or technical tools for the support of Data Governance has been selected for the interview process. The reason for this decision was that we expect that they would rather use the interview as a chance for advertising their products instead of focusing on the scientific aspects.

In fact, the selected interview partners are experts in their specific field, being it a concrete Data Governance job title or strongly related to one specific dimension of the DMBOK. By following this qualitative approach, experts from different organizations have been addressed in order to gain a broad insight into the topic.

Eight of the interviews were semi-structured interviews, whereas one has been a group interview with two people being interviewed since they combined their knowledge in their specific fields [26]. All interview partners have received the full list of possible questions in advance in order to be well-prepared for the interview. Nevertheless, some new appropriate questions have been discussed as well during the interviews. Each interview started with a short preliminary talk before continuing with the stepwise processing of the relevant questions within the given amount of time. In general, the

interviews with the chosen companies and institutions were conducted with essential decision makers and their technology departments. The following table depicts some information about the separate interviews:

Table 2: Details regarding the conduction of the nine interviews.

Date	**Duration**	**Job title**	**Type of Organization**
Sept 1, 2017	51 min	Data Governance Coordinator	Public organization
Sept 6, 2017	62 min	Head of Application Development	Public organization
Sept 6, 2017	39 min	Employee in Strategic IT	Large company
Sept 8, 2017	78 min	Team Leader Data Governance & Information Provision	Public organization
Sept 11, 2017	48 min	Market Risk Pricing Specialist	Large company
Sept 11, 2017	47 min	Chief Data Scientist	SME
Sept 14, 2017	77 min	Head of IT and Compliance; Data Privacy Officer	Large company
Sept 14, 2017	56 min	Software Developer	SME
Sept 20, 2017	43 min	Chief Executive Officer	SME

During the analytical part, the relevant information from the interviews has been put in relation to each other. Thereby, their content was used anonymously as well as formulated in a neutral and comparable way.

1.2.4 Online questionnaire

The third major part of the research approach and the second empirical fieldwork is an online survey which has been created upon the information gained from the literature review and the analysis of the interviews. The reason for its conduction is to check whether the information resulting from the interview process can be observed as well on a more general level.

The preparation evolved after the final interview and in parallel to the interview analysis. It was designed with the freely available online tool Google Forms since it is commonly known and relatively understandable for survey participants.

For the first version, it was used as a pre-test by asking one expert from a large company and another one from a public organization for further improvements. The final version of the online questionnaire consists of 33 questions (from which 24 are single-choice, 3 multiple choice, and 6 open questions).

Each survey participant had to go through all questions. The first questions focused on the type of organization and the industry [5], which is followed by a definition for Data Governance. In the main part, specific aspects of Data Governance are covered. In order to make the results comparable, eight of the single-choice questions have been scaled by using a 10-point Likert scale, from which two also have the 0 option for e.g. not knowing the concept, with

- "1" being a very low challenge (or opportunity) and
- "10" being a very high challenge (or opportunity).

Consequently, it was up to the respondent to evaluate the current situation of their organization. Moreover, open text fields were added at the end of these tables in order to present even more challenges or opportunities which are seen as such by the organizations. The survey was conducted entirely in English. In the scarce case of receiving a German answer, the comment was translated into English.

An invitation to participate in the survey was sent by email to all nine interview partners as well as to further organizations in German-speaking countries. These contacts were also encouraged to forward this survey to some of their appropriate contacts.

The survey started on September 25, 2017 and was open until October 14, 2017. Within this timeframe, 42 out of 44 participants have completed the whole survey. As for the organizational type, the distribution is given as followed: large companies (47.7%), SMEs (43.2%) and public organizations (9.1%). Most of the participants work in the industry of financial services and activities (48.8%), other industry (mainly software producers) not mentioned as a separate category (23.3%) and the information and communication industry (14.0%). Despite all efforts, the response rate for the public organizations was relatively low. One potential reason for this might be that these organizations did not deliver a response due to their limited knowledge in the area of Data Governance.

1.3 Limitations of research methodology

In order to study social or cultural phenomena within organizations, qualitative methods such as interviews are more suitable than quantitative ones since they focus on the meaning of phenomena. Therefore, textual data was mainly used where the information content is largely lost when they are quantified [36]. Although this research method provides an in-depth and first-hand insight into a real world setting, the results are constrained by the following limitations.

As for the interviews, a particular limiter is that they reflect an artificial situation, and some interview partners might not answer a question completely objective. That is, it could be the case that two employees from the same organization may be likely to provide different answers to the same question. Moreover, it is not possible to answer "all" questions within a single interview. A further main limiting factor is the number of cases that has been taken into account. In fact, the results may vary depending on the amount and type of organizations considered for this research, and are therefore not representative enough to conclude general assumptions.

In total, nine cases have been observed during the research process in the interviews. It should be mentioned that since the companies and institutions observed belong to different industries, it is not possible to generalize the findings by stating that the identified phenomena are applicable to every other organization of this type. In addition, especially for the SME sector, mainly IT-related (i.e. Semantic Web oriented) firms have been taken into account, which explains the relatively good knowledge and implementation state of Data Governance within these firms. Therefore, the findings cannot be stated to be applicable for most of the other (especially non-technical specialized) SMEs in Austria or Germany. Another limitation is that the findings from the interview analysis are focused on the current situation in Austria and therefore might not be applicable to other countries.

The same is also true for the online survey which could have delivered different results with more and other participants. The online survey had the goal to evaluate the awareness of the identified challenges and opportunities and to compare them with the findings from the interview analysis. Their comparison might be, however, problematic since the populations of the interview analysis and the online survey differentiate from each other. That is, the nine interview cases were all Austrian organizations, whereas the online survey has reached 78.6% participants from Germany, and only 16.7% in Austria. Consequently, the conclusions made from the direct comparison of these two research approaches may not be objective enough and are therefore only a description of potential phenomena identified within this

special survey.

In addition, the method of data triangulation can improve the validity and reliability of the survey outcomes. Although 44 different organizations were investigated, the results also remain rather with a lack of objectivity due to the limited number of cases observed. Since markets and companies and institutions are unique, there seems to be no suitable research method to depict the whole situation for all organizations in the respective field.

In total, it has been tried to keep the limitations as low as possible, especially by selecting appropriate and well-informed interview partners or survey participants. Nevertheless, each of the limiting factors remains, at least at a rather low level.

2 Results of the literature review

2.1 Definitions of Data Governance

Before defining Data Governance, we will consider the separate elements of its term:

"Data" can be seen as a set of characters which is of no meaning unless it is regarded in the context of a specific type of usage which turns these data into information [1]. According to Brous et al. (2016), the terms "data" and "information" are used interchangeably in many scientific sources, so that Data Governance can be also indicated as Information Governance. However, for this thesis, we will distinguish between these two terms since only the latter evolves from Data Governance. Despite the fact that the data value of the data assets is not represented on the organization's balance sheet, they represent an essential aspect since they are part of every internal or external activity. The six key data attributes are accessibility, availability, quality, consistency, auditability and security. In case the existing IT infrastructure is not able to fulfill these requirements, pilot projects should be used as a starting point for a new and reliable data integration system [10].

"Governance" has to be clearly separated from the term "management". The first can be seen as a high-level definition of necessary decisions in order to provide an organization with both an effective management and use of its IT in accordance with its overall strategy. On the other hand, management rather focuses on the concrete implementation of these decisions [40]. The following figure depicts fundamental concepts and dependencies in the context of Data Governance:

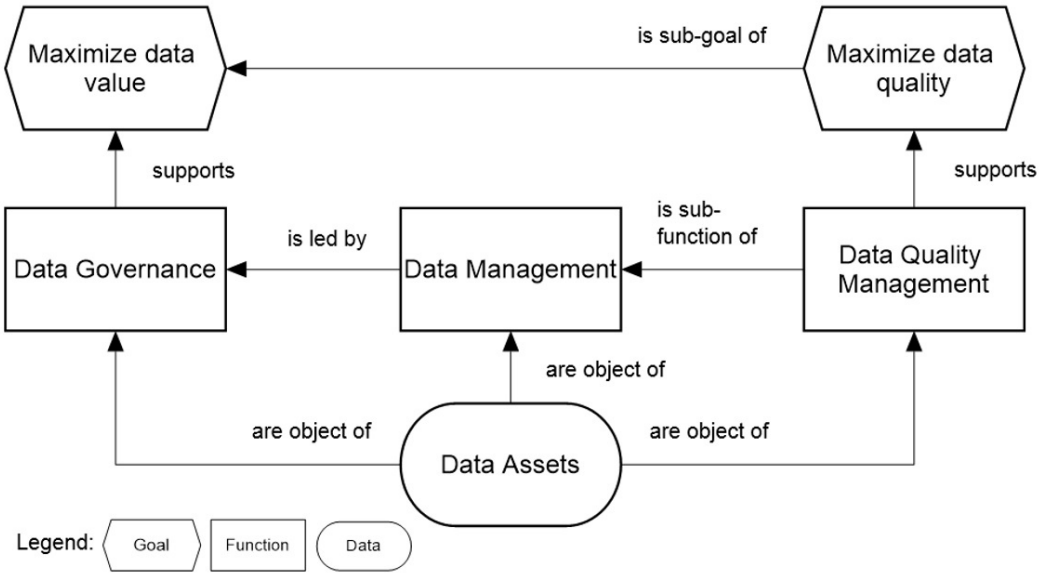


Figure 1: Fundamental concepts of Data Governance, data management and data quality management, based on Otto (2011) [53].

One important goal of Data Governance is to maximize the company's value of its data assets. An essential question being discussed until today is how this data value can best be determined. Within this context, data may only generate a business value if it is being used, which is directly related to maximizing data quality, being a sub-function of Data Governance [51].

Data Governance is used for the measurement and monitoring of different organizational aspects. Clear definitions of the data assets in each department are part of the data management process, whereas their integration into the organization belongs to Data Governance.

Due to the huge number of different organizations with their individual cultures, numerous possibilities to implement Data Governance exist. Consequently, Data Governance can be defined in various ways so there is no unique view which applies equally to all organizations. For this thesis, a variety of different definitions for this term from the existing literature have been considered and analyzed, and the detailing table can be found in the appendix. This process has led to the following definition:

"Data Governance is the definition of methodological, conceptual, organizational and technical rules, responsibilities, standards and procedures aligned with the organization's strategy and culture, with the goal of using data with their maximum potential within the organizational business processes."

This definition thus emphasizes the multidimensional meaning of Data Governance as well as its complexity.

2.2 The Data Management Body of Knowledge

In order to successfully implement a Data Governance strategy, various Data Governance frameworks are proposed within the existing American literature. The first edition of the selected DMBOK was published in 2009 and served as a comprehensive guide for best practices in the field of Data Governance. The second version came out in 2012 and expanded the scope by security issues, cloud services and ways to cope with an overall increase in the growth of data. Table 7 in the appendix depicts the key concepts of each dimension, thus giving an overview over this framework which covers 430 pages in total.

Whereas the first version focused on ten data management functions, the second edition identifies 11 so-called "Data Management Knowledge Areas" [20]:

1. *Data Quality*: defining, monitoring and maintaining data integrity and data quality
2. *Metadata*: collecting, categorizing, integrating and delivering metadata
3. *Data Warehousing and Business Intelligence*: managing analytical data processing and guaranteeing access to decision support data
4. *Reference and Master Data*: managing and reducing data redundancy, standardized definition and use of valuable data
5. *Documents and Content*: storing, protecting and enabling access to data in unstructured sources for integration and interoperability
6. *Data Architecture*: entire structure of data resources as an integral part of the enterprise architecture
7. *Data Modeling and Design*: analysis, design, building and testing of data models
8. *Data Storage and Operations*: structured storage and management of physical data assets
9. *Data Security*: guarantee privacy, confidentiality and appropriate access

10. *Data Integration and Interoperability*: acquisition, extraction, transformation, replication and virtualization

11. *Data Governance*

These dimensions support the identification of guiding principles for data management and also provide strategies for their maturity analysis or common organizational and cultural issues. The purpose of this framework is to offer organizations a structure for data management and therefore to enable their management of recognizing their valuable data assets [20]. The DMBOK Functional Framework maps seven environmental elements to each dimension. The basic environmental elements are defined through goals and principles, activities and deliverables. The supporting environmental elements (roles and responsibilities, practices and procedures, technology, organization and culture) are less structured than the basic elements. In addition, each of these elements has one type descriptor, which can be either people, process or technology [4]. A visualization of this relation can be found in the appendix.

Organizations can use this framework to establish their Data Governance strategy. In each chapter of the framework, it focuses in detail on each dimension separately, presents data management maturity models as well as giving an answer to further data management topics [20]. Finally, it should be further noted that at least the following two relevant topics are not included in this model, which are "assigning value to data" and "auditing data management" [28].

2.3 Concepts of data storage and operations

Four different types of data storage and operations will be taken into account since they could support effective Data Governance. The first concept is a so-called Data Lake, which represents a company's huge collection of all data sources or data sets in their raw format. It can be seen as a storage repository that uses dynamic analytical applications, with a massive scalability that can hold a huge amount of raw data in their native format and is stored until it is used for further operations. It also describes the processing systems which receive the raw data and keep the original data structure [43]. Therefore, various data formats can be found in this repository, with every element having a unique identifier and metadata tags. One advantage is that it can cope with vast volumes of unstructured data that arrive very fast in the organization, and are accessible immediately after their creation. So, the idea of a Data Lake is to derive further insights from these data. In contrast,

pre-built static data warehouses are established to store highly structured and slowly changing data [46].

In contrast to Data Lakes, Data Spaces are an organization-wide concept which model the relationship of different data repositories, thus connecting a variety of data formats stored in different locations [27]. However the concept of Data Spaces exists longer than Data Lakes, there is not much literature available which focuses on it.

In contrast, the concept of Linked Data is to create structured data that can be integrated and interchanged within an organization. It follows the idea of standard web technologies, which is to connect and query different data from various sources [8]. Enterprise Linked Data describes the technologies and principles that allow enterprises to integrate their data in a flexible way, by following a bottom-up approach [83]. This can result in internal and external values, such as a more efficient integration of the supply chain or cooperative engineering. Consequently, companies can benefit from this approach as they are able to combine their data much easier and therefore gain more knowledge from them, by saving time and transaction costs at the same time [22].

With the help of Linked Data, organizations can have their data landscape better under control and are also aware of their processes and organizational attributes. Concrete results can be seen in the improvement of governing corporate data assets and increasing their quality. With the growth of the systems, the diverse relationships support the understanding of how the various data are related to different parties, systems or initiatives through modeling their complex relations via ontologies. An ontology-based linked data technique is used by DeStefano et al. (2016) in order to strengthen the awareness and quality of the data. Another way to represent Linked Data is with the Resource Description Framework (RDF), which is a flexible graph that can be adopted at different levels within an organization. Depending on the actual level, employees can access various levels of detail [24]. Its effectiveness has been evaluated in some experiments and cases. In fact, also a white paper states that taxonomies and classifications from the search world are useful to be implemented in business practice [38].

According to Soares (2012), Data Governance principles can also be applied on Big Data. Big Data is a commonly well-known term which can be defined as such a high amount of data that is not possible to be processed with the help of a single standard device [66]. Currently, companies of all sizes are faced with ever increasing amounts of data. In order to remain efficient and reliable, these data have to be governed, managed, analyzed and visualized in an appropriate way to take actions necessary for doing business successfully, for example with the support of Data Scientists. This advan-

tage is also outlined in several white papers [60], [69]. Moreover, Priebe and Markus (2015) have provided a methodology which gathers and structures different data requirements which are used for improving data-intensive projects and thus enabling Big Data Governance [59].

Another emerging trend is Cloud Data Governance, which challenges have been identified and grouped by a business, legal and technical dimension [2]. Moreover, a well implemented and well-developed Data Governance is a particular precondition for realizing open data [80].

2.4 Key concepts and principles of Data Governance

Data Governance is not a completely new concept. In fact, the approach has changed to perceiving data not only as valuable within applications but to specific users extending this value for the whole company, thus resulting in a cross-functional and cross-architectural view that challenges business and IT leaders equally. In this sense, Data Governance needs an appropriate long-term support and investment to be successful [38].

So far, the focus of the scientific community has been put on structuring and organizing Data Governance [52]. A detailed study by Brous et al. (2016) have identified numerous key concepts within the existing literature. These are divided into four upper principles: organization, alignment, compliance and common understanding. The full list of these dimensions is illustrated in the subsequent diagram on the following page:

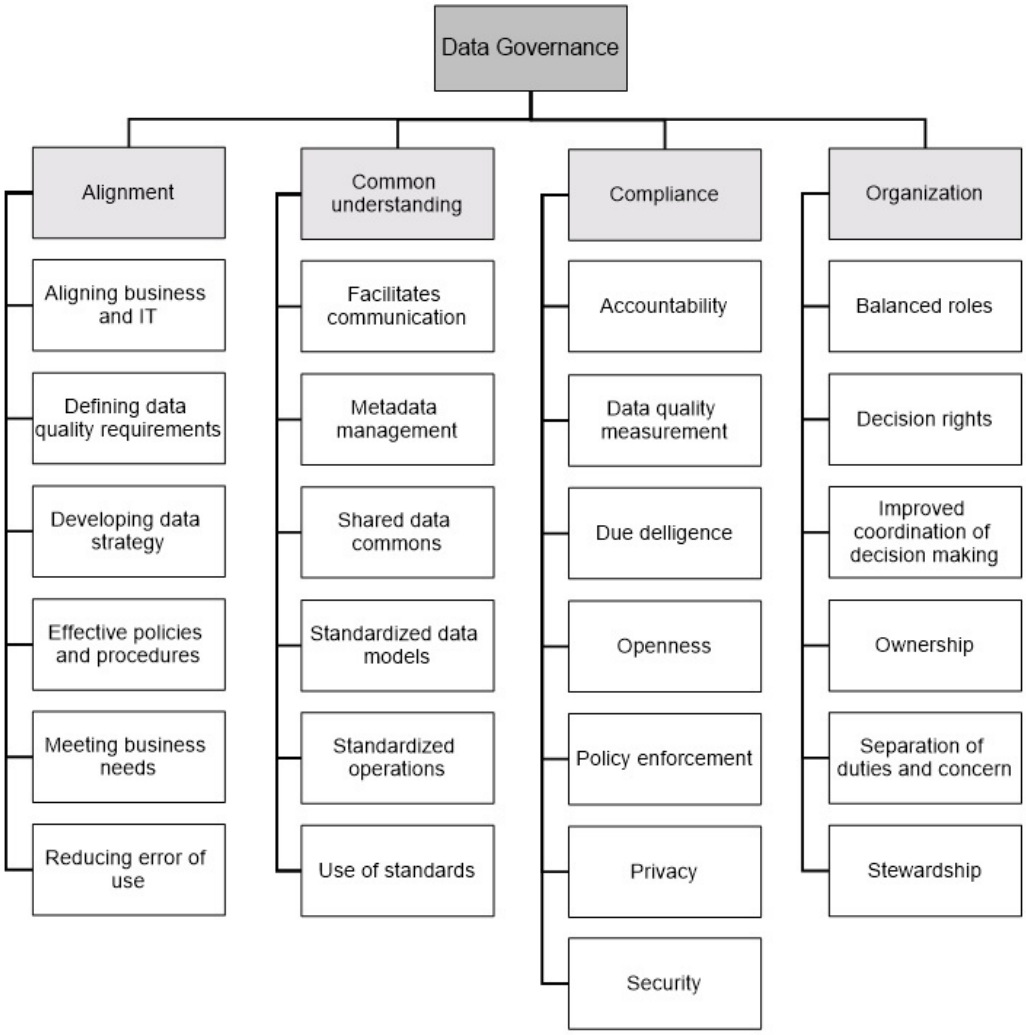


Figure 2: Long list of key concepts and principles of Data Governance, based on Brous et al. (2016) [10].

Consequently, Data Governance concepts and principles also imply data management to be compliant to strategical, tactical and operational policies of the organization. These findings can be an essential part for the development and implementation of effective strategies and approaches for Data Governance. Comparing to the DMBOK, some of its elements are included in this model. However, it can be stated that it focuses almost entirely on the organizational perspective. That is, it does not integrate further technical dimensions. Since Data Governance is widely regarded as a combination of both the IT and the business aspects, this model should be extended by technical concepts as well.

In case of the DMBOK, this mainly includes dimensions of

- data warehousing and business intelligence,
- data architecture,
- data storage and
- data integration and interoperability

in order to emphasize the strong relation between these two fields. In addition, an often cited phrase that should be considered in this context is "one size does not fit all", which means that it is not possible to create a unique Data Governance framework that can be used by all organizations without further adaptions [79].

Moreover, it should be said that this analysis of Data Governance key concepts and its four main principles mainly relates to public organizations. Nevertheless, Tallon (2013) argues that most of the inhibitors (i.e. challenges) follow an equal application across all industries whereas enabling factors can be seen as opportunities for organizations [68].

Researchers suggest that at least the following questions should be answered when it comes to Data Governance [40], [57]:

- What are definitions regarding corporate data that need to be made throughout the whole organization?
- Which roles does the decision-making process involve?
- How can these roles take part in this process?

Key aspects of Data Governance are people, technology [71] and processes [42]. In a study by Weber and Otto (2009), seven factors which are likely to influence Data Governance have been identified [78]. These are a company's strategy, organizational structure, competitive strategy, breadth of diversification, harmonization of processes, market regulation and decision-making. This model consists of three main components - data quality roles, decision areas and responsibilities - which create a responsibility assignment matrix. Therefore, it can be seen as a starting point when thinking about a Data Governance strategy.

Moreover, different levels of centralized, decentralized, and shared decision rights can be the right tool for managing diverse decision domains within one organization. A framework with five dimensions provided by Khatri and Brown (2010) can be used for the development of a Data Governance strategy

and approach for managing data as an organizational asset. This includes data principles (clarify role of data as an asset), data quality (establish requirements of targeted use of data), metadata (provide the semantics of data such that it is interpretable for its users), data access (specify access requirements) and data lifecycle (determine the definition, production, retention and retirement of data). For each decision area, data is the driving factor and therefore strongly supports the company's corporate strategy [40], [12]. Each of them is identified by specific domain decisions and also names potential roles. In comparison with the DMBOK, it is more intuitive and rather focuses on treating data as an asset and as a business function. In this context, especially the data lifecycle approach supports this thinking [30]. Further decision areas by Pierce et al. (2008) include company-wide standardization of data definitions, logical data mode and standardized business rules [57].

Within the community, similar core areas of Data Governance have been formulated in white papers. One example is the one of IBM (2007), which also regards data quality management, information lifecycle management and data protection, and additionally sees the supportive tasks of data architecture management, metadata management and documentation of review results. IBM has also created a maturity model for Data Governance, which includes eleven different categories: organizational structure and awareness, stewardship, policy, value creation, data risk management and compliance, information security and privacy, data architecture, data quality management, classification and metadata, information lifecycle management, audit information and logging and reporting [19]. This extensive model can be used by organizations with experience in Data Governance to improve their related policy.

If we look at Data Governance, it is linked to a certain extent to IT Governance since data management is seen as a discipline of IT management. Both concepts are considered as being part of a company's corporate governance. Organizational issues that are not within the scope of IT management are part of data quality management. Therefore, Data Governance defines all necessary decision rights, accountabilities, standards, rules and policies for a subsequent data management [82]. These ideas are summarized in the following graphic:

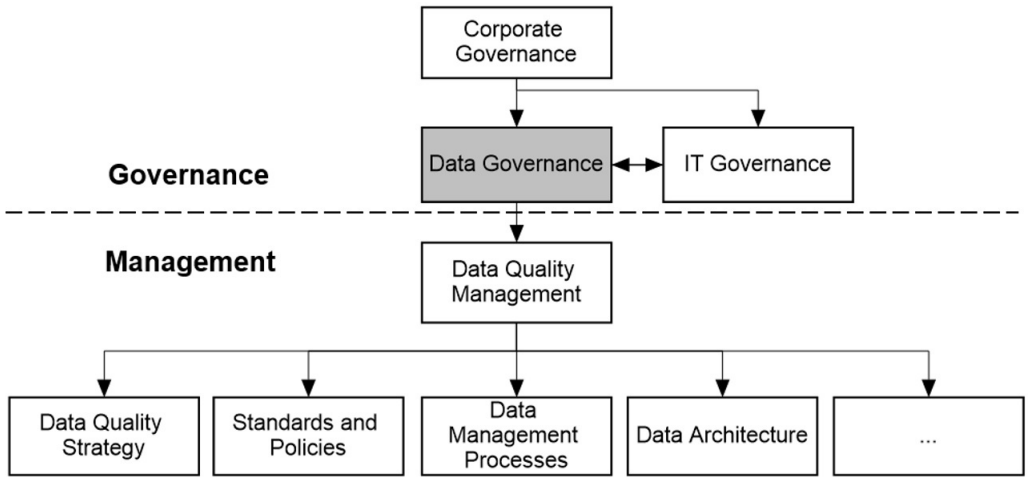


Figure 3: Relationship between Data Governance and data management, based on Wende and Otto (2007) [82].

Data need to be managed in an appropriate way to realize their full value-creating potential by ensuring that they are competitive and agile so that the needs of the customers can be met [51]. One approach for using data efficiently is to combine them with metadata [50]. Besides being effective, Data Governance should also be transparent and auditable. This transparency can be reached by using a multi-level metadata management approach [75]. In fact, the integration of the data from the metadata repositories can contribute a lot to this process. Data integration combines business and technical processes in order to ensure that data from a variety of sources are combined to meaningful and valuable information assets for the organization [41].

A Data Governance program also supports the implementation of master data management systems which helps to ensure that heterogeneous databases and applications are integrated and interoperable in a business oriented manner [6], [75], which relates to the concept of Enterprise Linked Data.

The question of an integrated view on data goes hand in hand with the corporate-wide accountabilities for data quality management, which comprises professionals from both business and IT departments. Nowadays, numerous companies still seem to strictly assign the accountabilities and responsibilities for data quality management mostly within the IT departments and thus ignoring organizational issues.

One potential problem is the IT budgets of these companies and organizations. Research by Tallon (2013) in this area has shown that 47 percent of the IT budgets are assigned to maintain the IT infrastructure, 40 percent goes to information and transaction processing, and the remaining 13 per-

cent to strategic IT investments. In the closing years of the 20th century, storage issues and related hardware and software tools were mainly seen as a company's competitive advantage. This view changed with the upcoming of increasing amounts of data, thus making data a more strategic tool - and software a more tactical tool [68].

Without an accurate implementation of Data Governance, companies will be vulnerable to significant risk of failure in providing an effective compliance and corporate governance. Decreasing value of its data assets derives directly from a wrong execution, which can even result in a fine due to liability issues [30].

As data is ubiquitous in a company, Data Governance should be integrated horizontally. One main task is to assimilate data throughout the whole company and to support individuals in overseeing the administration of data-related processes. That is, Data Governance also involves concepts of metadata, unstructured data, registries, taxonomies and ontologies. When it comes to terms of more complex big data integration, which can be seen as the combination of technical and business processes mainly used for the combination of data from diverse sources into meaningful and valuable information, their governance becomes even more important [41]. Last but not least, a well-developed Data Governance and clear languages applied to guidelines for sharing data in the private and public sector are one of the preconditions needed for open data in interoperable repositories [80].

2.5 Implementation of Data Governance

Data Governance can be seen as the starting point for an effective data management program [41]. Moreover, the success of Data Governance strongly depends on clear definition concerning the objectives, metrics and processes involved.

At the starting point of introducing a Data Governance program, the questions arises whether it should be driven by the IT or business department within an organization and what parts come to IT Governance. In the scope of the IT Governance lies the guarantee to align the IT infrastructure with the organization's business objectives in a cost-efficient way [44]. In general, the IT provides the infrastructure necessary for processes and reporting on the data. This infrastructure system has to ensure a variety of different aspects, such as data security and privacy. Besides, it is further stated that it should be the function of the business to coordinate a Data Governance program since it needs the data for appropriate decision-making. Consequently, the business departments and the IT need to collaborate [41].

The goals for Data Governance can be separated by formal business and

IS/IT-related goals as well as functional goals. The first include the assurance of compliance, enabling decision-making improvements of customer satisfaction, increase in operational efficiency and the support of business integration. The IS/IT-related goals are to increase data quality and to support IS integration. Functional goals are the creation of data strategy and policies, establishment of data quality controlling, of data stewardship, data lifecycle-management and data architecture or the implementation of data standards and metadata management [52]. Various levels of the organization are brought together and work towards specific goals, which enables the alignment of data-related programs with the corporate strategy.

Data Governance and its implementation are also seen as an ongoing, evolutionary process. As Data Governance is a comprehensive concept, it affects every employee and is definitely not limited to the IT department. If the data are left to the IT, it is not possible to manage them as a corporate-wide value since they cannot play a supportive function in business related activities. So the technology is the fundamental basis for governing, managing, monitoring and standardizing data as well as to reduce risk, improve data quality and offer a more secure data culture. Therefore, a critical success factor is that the senior management will need to realize the value of data, and establishes a corporate vision and a "positive data culture" within the whole organization [30], [62].

After having clarified the role of the IT, specific needs in relation to Data Governance should be understood and sorted by their relevance. As a next step, a concrete plan in combination with a specific framework is recommended for further proceedings, which reflect the business objectives and priorities as well as the organizational structure and culture [67]. As a next step following from a clear structure as well as related roles and responsibilities, suitable repository tools for data profiling and metadata are needed [15].

2.6 Best practices

Very few best practices for Data Governance are mentioned in the literature. Tallon (2013) divides them into structural, operational and relational practices [68]. However, these have not been created within the German-speaking countries and therefore might not be appropriate or easy to implement.

Suggested best practices include

- taking a holistic approach but starting small;
- ensuring executive sponsorship;

- defining data stewardship at an early stage;
- establishing quantifiable benefits by building business cases;
- establishing, collecting and reporting on metrics in order to measure the progress; and
- linking incentives in order to award participation [67], [25].

Rewards for appropriate Data Governance behavior can support this process effectively. When introducing it, Data Governance should be obliged and not an optional role as the organizational success depends on compliant practices. In addition, an effective Data Governance program needs time and it is necessary to find the right amount and right level of control for it. Maturity levels are one way to identify the current implementation phase. For example, a six-level Data Governance maturity model (with the milestones none, initial, managed, standardized, advanced, and optimized) is provided by Oracle (2011). It is stated that the focus of this model should be adapted to the growing scope of governance over time as well [67]. Another potential maturity model is offered by Gregory (2011) in order to implement a Data Governance strategy to improve the organization's data management capability [30].

In recent years, the correct management of data has become one of the key factors for a company's business success since it supports the maintenance of competitiveness as well as it allows to pro-actively fulfill its customers' needs and to keep costs under control. Companies and public organizations of all sizes and complexities attempt to reach an accurate management of their data as an essential asset that is shared and reused across numerous software applications and systems, business processes, departments and users throughout the entire organization. It could be stated that organizations need the establishment of detailed standards, policies, and processes for an accurate usage, development and management of data, which goes hand in hand with an accurate Data Governance model. In the light of this, the organizational structure and development of the technological infrastructure as a continuous support for the governance of their data is also required [56].

2.7 Data quality

The quality of data can be seen as an essential aspect of an effective Data Governance program. There are numerous outside influences that affect data quality. The governance of data is therefore dependable on a proper understanding of the data itself and their importance for the organization. One

potential and often cited data quality framework has been proposed by Wang et al. (1996) [77]. According to Brous et al. (2016), Data Governance can contribute to better data quality results for an improved decision-making process, such as in asset management [10]. This can also reduce related ethical, organizational or technical barriers within the organization and requires a constant quantification and measurement of their data quality [41], what implies that these data need to be governed.

Driving factors in data quality are availability and discoverability, trust and authenticity, acceptability, accuracy (comprising correctness and consistency), applicability, integrity, completeness, understandability and usability [18]. Several factors to measure the quality of data include timeliness, relevance, completeness, trustworthiness and contextual definition. Therefore, a suitable and effective data management program is needed and depends on a continuous work and support between the business operations and the IT [41].

Master data quality as a further concept in accordance with a Master Data Governance is only effective and successful when those policies and procedures support a business in running seamlessly. Their actual performance should be measured with the help of key performance indicators (KPIs) and will determine their level of consistency and compliance [60].

2.8 Organizational roles and responsibilities

It is stated by Weber and Otto (2009) that a corporate-wide data quality management is connected to the process of defining decision rights and accountabilities [79]. Moreover, Thomas (2006) argues that one particular problem when implementing Data Governance is the lack of authority and clearly distributed roles and responsibilities [70], [15]. The major challenge of having an appropriate Data Governance policy is the definition of an accepted and suitable Data Governance model with shared responsibilities. A supportive tool for the development of an individual role concept is using a RACI notation. Hereby, data quality roles, decision areas, main activities and responsibilities are precisely defined for employees throughout the whole company for designing and defining decision areas, authorities and roles [82], [79].

Given this context, an important step in order to create an effective Data Governance program is to define its structure. This ensures a basis for a transparent decision-making process. The definition of roles and responsibilities are made in a way that they are accountable for their actions [15]. It can be stated that, dependent on the source and type of data, different people hold the responsibility for the data and its quality. For example, financial

data is surveilled by the Chief Financial Officer. In the light of this separation of responsibilities, the business department is responsible for ensuring the correctness, availability, reliability and fitness of the data.

Within the existing literature, numerous recommendations for different role models of an effective Data Governance program can be found. Several organizational roles and committees which are responsible for the achievement of different Data Governance goals (formal and functional) exist. These include a master data management council, data owners, lead and technical stewards, steering committees, master data owners and officers, Data Governance managers, data quality managers, data stewardship managers, "data responsibles" or data architects. According to Otto (2011), the most commonly mentioned roles are data stewards, data owners and data committees [52]. As for the roles involved in Data Governance programs, the focus has been put on the subsequent roles which were also named within the interviews.

Data steward: has a data management function, works within a specific division of the organization and ensures that the business departments have the appropriate use of the data. Moreover, this role evaluates various data-related problems and has the detail knowledge of the IT, data requirements, and business processes for developing the data within their respective decision domain [79]. Other related roles are chief steward, business data steward, technical data steward [81], [38].

Data owner: is part of a specific business department or division and gives a detailed specification on the data-related business requirements and the data quality [40]. However, it does not mean that these employees "own" the data per se since this is true for the organization as a whole.

Data expert: is located between data stewards and data users and supports the data users in understanding and using the professional data quality requirements. Moreover, this role supports the technical design of the implementation of professional data quality tests [23].

Data users: are data stakeholders both from different departments or subdivisions within an organization. They can use the data on an operational level and also have the responsibility to report on any data related issues to make the use of the data more efficient [15]. Data Scientists are an example for Data Users.

Chief Data Officer (CDO): ensures the roles and processes to guarantee data principles, which also involves the exploration of new technologies and frameworks, relational and non-relational databases, for example to detect suitable physical repositories for each data lake repository [45].

Data Governance council: this council involves executives from different divisions within the organization who are accountable for managing data as

an asset. Their responsibilities include the endorsement of policies and alignment of the business and the IT on a strategic level [15]. A Data Governance Coordinator can be part of this entity, where the head is in many cases the Chief Information Officer (CIO). As being part of the board of an organization, a Chief Data Officer (CDO) exists on the same level as the CIO, especially in organizations which are very data-driven [34].

Data Governance steering committee: is the institutionalized central board for decision-making concerning Data Governance on a strategic level. Thereby, it takes into account the various interests and demand of the departments and business divisions within an organization (representatives are the data owners) and the data management function on the other side (representatives are the data stewards) [40].

Data Protection Officer: can be introduced to surveil data protection regulations, especially the General Data Protection Regulation (GDPR) and has to show expertise knowledge in data security and data protection practices within the organization [9].

Furthermore, organizations need Data Scientists in order to provide a new inside into datasets and thus generating new value out of them [74]. In essence, it is one idea of Data Governance to ensure that cross-functional teams are brought together (e.g. data stewards and Data Governance coordinators) in order to solve certain issues or to serve data stakeholders with specific services [41], [64].

2.9 Impacts of Data Governance

This section provides an overview over the current impacts of Data Governance within the three different organizational types, which have been identified within the existing literature. Then, further general challenges and opportunities of Data Governance are considered in more detail and are sorted by the following categories, which evolved during the research process: organizational and cultural, monetary, technical, strategy and business, legal, knowledge and experience, security and privacy, and data quality. These are then linked to the summarizing table in connection with the results from the interviews, which can be found in the appendix. Overall, it is stated that enablers of Data Governance can be related to a specific industry, whereas inhibitors would rather be linked to the same application across all industries [68].

2.9.1 Large companies

Few detailed investigations have been conducted for the situation of Data Governance in large companies. Otto (2011) investigated the situation of two large companies from the telecommunications industry in Germany and the UK. It is stated that there are numerous different options for a Data Governance program: one of the companies follows a bottom-up approach in combination with project-driven decisions, whereas the other focuses on an authoritative top-down philosophy [53].

A major reason for them for having Data Governance is to achieve an overall good data quality and to treat the data as a valuable business asset. This included the provision of huge amounts of resources and has also required a fundamental organizational change within these companies. However, changing organizational aspects of a company is a long process which sometimes takes several years to show visible results [53].

In a additional paper, Otto (2013) observed a single-case study of a large North-American consumer products company on how to measure the effectiveness of Data Governance by regarding it on a lifecycle curve. Thereby, it considers data quality management measures over time [54]. Another paper surveyed data management strategies in large agencies of the private and public sector to identify best practices [72]. Furthermore, in a recent study by Riggins and Klamm (2017), it is observed how a large public accounting and business consulting firm treats its data values with the help of Data Governance and how this interacts with Business Intelligence and Big Data [63].

Companies within the financial services and health care industries are main adopters of such a programs for apparent reasons. For organizations from the financial services industry, Data Governance strategies vary. Some firms may focus on governance bodies and councils, whereas others focus on data stewardships and workflows. In order to show the value of data as a tangible benefit, appropriate quantifiable key performance indicators (KPIs) and other metrics could be used. However, this remains rather difficult as continuous cost-benefits or ROIs are not easy to quantify. Since there exists no "one size fits all" model, various suitable IT implementations and software tools have to be carefully identified in coordination with the specific business goals of the organization [14].

An Europe-wide study with 45 banks has further identified Data Governance as the key to successful data management and to improve data quality in particular. However, current Data Governance frameworks lack clear responsibilities for managing the data. The survey also states that specific departments or functions are required for introducing Data Governance ef-

fectively and clearly [61].

2.9.2 SMEs

Academic research for the adoption of Data Governance within SMEs is relatively scarce. Begg and Caira (2012) have surveilled the situation in the west of Scotland by investigating ten different SMEs [7]. They surveyed the implementation of a small and efficient Data Governance framework proposed by Khatri and Brown (2010) because it was easy to use. Their findings include the fact that none of these enterprises has thought about introducing any kind of Data Governance in its primary understanding. In fact, they had their own data management strategies, but without following explicit rules throughout the whole enterprise. However, the adapted Data Governance approach has improved their business especially in terms of "having better data".

In fact, Data Governance frameworks are so far only poorly implemented for SMEs, and they are not likely to be adaptable and scalable in an appropriate manner. Furthermore, the knowledge required for handling and understanding data-based issues and technology is often not present in a small enterprise. Another problem of this sector is that the classification of roles and responsibilities for such an approach has not been clearly defined so far [15].

None of the Data Governance frameworks published so far seems to be suitable for being used within the SME sector without severe adaptions. Although a number of these frameworks claim being adaptable and scalable to their specific needs, this has not proven correctly. Without advanced or even basic technical experience or enough specialized personnel, these frameworks are also hard to understand. Since the monetary possibilities are in general limited, SMEs face difficulties in funding Data Governance models.

The effort necessary to implement Data Governance is regarded in many cases as far higher than the resulting benefits. Moreover, the acceptance of an upper Data Governance program is not given before a effective IT system guarantee is provided by IT governance. In addition, SMEs sometimes fell like being treated as a "sink for data" since amounts of data from different external sources (e.g. customers or suppliers) flow into their organization. Ineffective and inefficient data management can result from being a mere follower in relation to data format and use. Another problem can be that SMEs are not able to identify the inherent value of their data, especially for their business and independent from the technical systems.

The adaption of specific internal systems in order to deal with data of various forms from external partners also needs to be considered. In addi-

tion, a potential lack of technical experience in SMEs could prevent these organizations from easily adapting their existing IT systems to new business changes or practices.

It should also be considered that the criteria which distinguish SMEs from large companies are no necessary factor for determining the data environment of these organizations. Therefore, it could be the case that for example a medium-sized enterprise has much less data flows (in terms of volume or nature) than a small enterprise which is focused on a data-driven key business. Moreover, it can be the case that due to the growing appearance of technological advances (e.g. cloud computing), SMEs make a greater use of digital data. It is also stated that even if an SME does not follow a Data Governance approach, third-party cloud computing solution providers might already work with such a data strategy and therefore enforce its clients to do the same (e.g. in terms of data interoperability).

The SME sector plays an essential role in the worldwide economy. With Data Governance, these enterprises should be able to identify their data values, what would strengthen their competitiveness in an increasingly digital business environment. Therefore, by using e-business systems or extending their business strategy with a stronger use of IT systems, SMEs will collect huge volumes of data which have to be governed accordingly through Data Governance. In addition, it is stated that prospective regulatory requirements can force SMEs to work with data-related issues [73], [40]. Especially on SMEs in the banking and financial services industry, these regulations already have a major impact and change their business models. It is also proposed that this growing trend would appear in a worldwide context, which makes it even more beneficial to implement a Data Governance program [7].

2.9.3 Public organizations

Large volumes of data about individuals are collected and permanently stored within various public organizations. One particular opportunity is to analyze these data for the benefit of society and another to support these members in improving their decisions [10]. However, it often is the case that a nonexistence of Data Governance and/or related implementation errors can restrict this application. As Thompson et al. (2015) have shown, public organizations often lack methods, effective tools and the knowledge to handle these amounts of data. Their study illustrates that these inabilities and inefficiencies are due to a missing Data Governance policy and not the result of the existing business rules or technological aspects [72].

Due to their complexity and scope, Data Governance projects have often failed in the past within government organizations. Moreover, no common

Data Governance framework exists for the implementation within public organizations, so concrete decisions should be defined for each organization individually. Another inhibitor is that most of the standards, polices and mandates concerning Data Governance currently available need a lot of time to be analyzed and understood, what is also true for ensuring a common understanding by the majority of public sector workers.

Four formal Data Governance goals for public organizations have been defined by Otto and Weber (2015): enabling better decision making, ensuring compliance, increasing business efficiency and effectiveness and supporting business integration [55]. Managing their infrastructure assets is crucial and challenging, and has to be compliant with the area of new data resulting from social media or the Internet of Things. Potential requirements are outlined by a survey of [11].

In another case, it was looked at the potential of building a metadata governance model within federal government departments and agencies in the US and show their enabling factor for information sharing [16]. A further paper investigates the impacts of putting public data and operations into the cloud. These are relatively common in the sector of public institutions since they want to outsource the risk of their data. However, a number of access-related and security-based problems occur. Since these public sector data are important and thus need to be governed accordingly, public organizations may gain specific rights for using cloud solutions or even new cloud concepts will be created [29].

Another issue is data security, focusing in particular on privacy and data protection. In a white paper, it is stated that public organizations should maintain and extend these measurements for the data they manage, such as intellectual property, market data and personal information of internal and external customers. In the light of a rise in compliance obligations, such requirements are needed to be fulfilled as well. With the help of Data Governance, these aspects can be approached but it needs a cross-disciplinary effort by considering different dimensions (e.g. information technology, human resources or legal, business units) [65].

2.9.4 General challenges

2.9.4.1 Organizational and cultural

The organizational structure in general has been identified as major inhibitors to have an effective Data Governance policy [68]. In many cases, it is rather the lack of structure than the lack of will. Cultural barriers are also named as a particular reason for potential failure of Data Governance

projects [47]. For example, Data Governance should be considered to concern the organization as a whole and is not "just another IT problem" [30].

Several policy-related issues have to be addressed, such as data stewardship principles or effective data management approaches in order to realize the full data potential [35]. Individual Data Governance configurations are unique and define roles, decision areas and responsibilities. Moreover, people with appropriate specializations need to be hired, trained and integrated into the organization. Therefore, first frameworks have been developed [40].

Potential problems include the lack of having defined clear roles and responsibilities and also the missing mandate to carry out data quality improvement initiatives. A clear definition of responsibilities for the various data (data ownerships) is needed and helps to avoid redundant data or data reconciliation and simultaneously eases reporting processes.

Another challenge is when an organization lacks the role of a CDO whose responsibility is to coordinate Data Governance strategically at the same level as a CIO. Moreover, the introduction of a Chief Information Security Officer is rather controversial, but can result in improved work with regulatory requirements, for instance in the banking sector [61].

2.9.4.2 Strategic and business

During the last few years, the awareness of organizations about treating "data as an asset" has increased. Another challenge is to focus on the identification of fundamental decisions that need to be made and by whom they should be made instead of looking separately at day-to-day decision making [40]. In general, the alignment of the data with the business needs should be ensured, especially in terms of qualitative data ("fitness for use") [52], [53]. However, this alignment needs to be initiated and governed by the top management and therefore should exhibit a measurable success [47].

Moreover, the creation of a Data Governance policy which comprises all eventualities or possible uses for data is rather difficult. With focusing too much on the technical and organizational risk involved, important data might not be used to their full extent and does not result in full value creation. A too complex mix of products and services as well as strategic misalignment is also mentioned [68]. Another potential limit is that beyond a certain point it might result in counterproductive effects when continuing with Data Governance, and thus limiting the value which can be obtained. For example, this could be the case if an employee is not able to perform specific task because he encounters other Data Governance limitations [33].

In addition, data-related problems in relation to operational efficiency

include 1) difficult negotiation with stakeholder due to ineffective data exchange between these two groups, and 2) long data processing durations or wrong and incomplete data from external sources. Therefore, an appropriate Data Governance framework ensures that the data are as complete and accurate as possible so that they can be utilized in the systems of third parties by defining the standards for interacting and collaborating with them [56].

In a white paper, further challenges are described. At first, it is important to have a clear concept and a realistic plan. Since many organizations are not sure about when and where to start with the implementation, they might not do anything ("analysis of paralysis"). A concrete analysis of the current and the desired prospective state with the help of a maturity model for a clear time frame can support this process effectively. As a next step, some of the specific dimensions of Data Governance mentioned above should be taken into account as well [58]. It is further stated in another white paper that the failure of many Data Governance programs results from the complexity and difficulty of implementing such an organizational-wide program and that it is hard to determine its value. Its success is dependent on following a structured and agile approach in connection with strong leadership and management activities, which can be hard to achieve [32].

2.9.4.3 Monetary

In addition, costs are a main issue and should therefore be analyzed especially under the aspect which parts of an organization take part in this continuous funding as this decision can influence the question "where will Data Governance live". Further benefits include increased revenues, compliance improvements or savings in costs. Another problem is to produce a ROI for Data Governance. One monetary challenge is to see that Data Governance is not only an extra cost factor of the business. This also involves a concrete estimation of the costs involved in such a project, which is an inhibitor by itself [30].

2.9.4.4 Technical

In many organizations, core business-process systems are relatively old and do not operate in accordance with clearly defined roles and responsibilities for data ownership. These often do not work well enough with Data Governance, with the result that their data are stored in data silos. This also harms data analytical tools, and is intensified in case of a decentralized IT. Consequently, the efficiency of the related business processes decreases

and adversely influences decision-making. Another problem in this relation is the bureaucracy which surrounds these data and which reduces the ability of adapting strategies in large organizations [60].

The increasing amount of data, their lack of quality and appropriate integration are the key challenges for a Data Governance program. It is also regarded as being difficult to solve the problem of the data spread and to reduce data silos. A further technical challenge which should be faced by Data Governance are a packrat culture (data hoarding) and legacy IT systems (weak integration) [68]. It is also a challenge to ensure an appropriate data architecture for having a flexible approach for data processing and data delivery [21].

In a study by Cheong and Chang (2007), it is also mentioned as a problem that the data assets may not be easily accessed. Moreover, it could be the case that data standards are difficult to define on a technical level as well or to reach consensus among various departments [15].

2.9.4.5 Security and privacy

The appointment of a Data Protection Officer or a Data Privacy Officer as Data Governance responsibilities for these two areas at the right level is also regarded as challenging [21]. Moreover, the secure storage of sensitive and strategic data (e.g. business plan, sales figures, financial and production metrics) has to be provided [47]. It is further stated in a white paper that Data Governance strategies should also include high availability and disaster recovery, what can be difficult to achieve [69].

2.9.4.6 Data quality

Another potential challenge is to have too high targets for improved data quality. For its measurement, appropriate key performance indicators (KPIs) should be developed and used accordingly. Poor Data Governance can result in the risk of data theft and loss, breaches of data privacy and or regulatory compliance and damage from poor data quality [30]. Without the creation of a fundamental understanding of Data Governance and the awareness of the connection to its own corporate data landscape, governing data and its resultant data quality may not be achieved.

2.9.4.7 Legal

Regulatory and compliance requirements are a major challenge that could be faced with Data Governance [38]. For example, a recent survey in the banking sector has shown that the main need for Data Governance is the fulfillment of regulatory requirements and corporate governance policies [61].Such regulations can vary by region (e.g. EU or the US) and on the other hand be influenced by the absence of industry-wide data standards [68]. Thus, another open challenge is how to define appropriate legal rules for the Data Governance in the age of digital economy. This may only be reached with policy decision-makers and therefore increases the scope and meaning to this external perspective [39].

2.9.4.8 Knowledge and experience

A fundamental challenge is to create the understanding of the usefulness of Data Governance among all responsibilities within an organization. Therefore, Data Governance policies might not be conveyed well enough and are seen as unnecessary or excessive, which can for example result in not completing the assignments [33]. In fact, Data Governance guidelines and policies require a further analysis and understanding before being implemented [82], [10]. Another challenge derives from a lack of understanding that business definitions vary and therefore underestimating the amount of work involved in Data Governance projects and trying to move too fast from no-Data Governance to enterprise-wide Data Governance [47].

2.9.5 General opportunities

2.9.5.1 Organizational and cultural

Organizational success factors include a clear definition of roles and responsibilities, a collaboration of the business and the IT, executive sponsorship by senior managers, and the creation of an integration competency center (also: center of excellence) to support the underlying data integration technology [56]. That is, a relatively high alignment of business and IT facilitates a Data Governance program due to the closer cooperation between responsible authorities, for example between the CIO and business executives. With the help of Data Governance, a centralized IT and organization structure can be achieved [68]. A Data Governance framework can be further used for marketing issues of the topic inside the organization [66]. It is also

needed to communicate its relevance in a common understandable way and to ensure compliance of internal governance policies [56].

2.9.5.2 Strategic and business

Data Governance can help responsible employees to make better decisions and to consider and represent the needs of their internal and external stakeholders. A reduction in operational friction is also possible and it supports a broad understanding of data issues within the organization. Moreover, an increase in effective and more transparent processes and the creation of repeatable standard processes are enabled [37], [41].

A holistic and unified approach and view on the data will result in more profitable strategies. Such a strategy comprises the main elements people (executive sponsorship), policies (accepted guidelines for collecting and managing the data) and technology (data integration tools and appropriate data models) [21].

Main strategic benefits resulting from (Big) Data Governance are that the data can be found, described and managed in a very effective and reliable way [69]. Further business drivers have been identified in a white paper and include improving corporate flexibility, business agility and reducing business risk [49]. A highly focused business strategy, and the promotion of the strategic use of the IT are possible [65]. It can also deliver a predictable rate of data growth which can result in multidimensional benefits [68].

In addition, it can also lead to overall improvements in efficiency [31], [35], [42], to increasing revenues and market share [7], to perceiving how information initiatives can perform [31], [42], to the common trust in information products [53], and to accepting monetary expenses on information management projects [56], [68], [72]. Another point is to maintain business ownership of information assets within the organization [3]. It can also promote well-informed real-time decision making, improve customer service quality and operational effectiveness [14].

2.9.5.3 Monetary

At first, it can be stated that a Data Governance program can add value by a consistent view of the information enabling the IT to enter a more strategic level instead of only focusing on supportive tasks. This improves efficiency by granting access to the relevant information assets, what can result in improved planning and cost savings on an operational level [3]. Business drivers, which can be also seen as opportunities, are to grow revenue,

to lower costs and therefore to increase the organization's profitability [56], [14].

The general aspect of cost reduction has been argued in a number of papers [56], [53], [54], so the implementation of related data management issues can be cheaper and more effective [61]. Besides, also a white paper discusses a relatively cost-effective approach and related guidelines, which also delivers measurable improvements within a short period of time [13].

2.9.5.4 Technical

Data Governance can also be used for coordinating big data analytics and supporting the automated processing of high volumes of data. It is therefore used to strengthen the trust in qualitative information that resulted from a Big Data analytical method and to convince all related people involved in the process of gained information from huge amounts of data [66].

An appropriate Data Governance program can also ensure that the enterprise data is not inconsistent, unreliable and unrepeatable. Since departments are rather focused on the fulfillment of their own data-related needs and requirements, inconsistencies and data redundancies are probable to evolve. Therefore, it is needed to get the best out of the data by taking the right steps in time [21]. A potential solution to ensure scalability and consistency of a Data Governance program is the provision of a unified enterprise data integration platform, which particularly grants access to all organizational data, offers a platform-neutral architecture within a heterogeneous IT domain, and simplifies the overall data integration lifecycle [56].

In addition, the following technological success factors have been identified: access to all data independently from their source or structure (e.g. from relational databases, XML or messages), discovering poorly documented or unknown sources, cleansing the data to ensure their quality, integrating them by having a consistent view across all systems and combining fragmented information from different systems, delivering data in their right format and at the right time to each user, developing and managing the reliability, scalability and performance of critical enterprise systems and auditing, monitoring and reporting of the data to identify their improvements over time [56]. Further technical benefits deriving from a Data Governance program include operational business intelligence initiatives, enterprise data warehousing and traditional business intelligence, information-as-a-service or operational software updates [38]. A further aspect to consider is that data corruption as a result of user error could be reduced by the capability to create snapshots and tracking data histories in combination with data recovery options [69].

2.9.5.5 Security and privacy

A pivotal role is privacy [66], [3], [73]. A major benefit can be the result of a reduction of privacy violations [68], [72], [73], an increase in data security [40], [56], [72] or a reduction of the risk of regulatory and civil liability [54], [68], [14]. Moreover, Data Governance can also support the identification of different kinds of risk (e.g. acceptable vs. affordable risk) [68].

2.9.5.6 Data quality

Another possible opportunity is the improvement and maintenance of data quality in order to meet strategic business relevant requirements (e.g. compliance issues or integrated customer management) [52]. Data Governance can also motivate the use of data since data are only creating value when they are used within the organization [68]. In strong relation to business intelligence, it helps to achieve data validity by ensuring data integration [3]. In many cases, the establishment of Data Governance becomes a major issue when the management starts to discover flaws in the quality of its data. Moreover, data quality and data integration can reduce the costs for data cleansing [49].

2.9.5.7 Legal

As for the legal aspects, it is the goal to meet regulatory or compliance requirements [68], [66], [56]. Another legal driving factor for the need of Data Governance can also be mergers and acquisitions [38].

3 Results of the survey

3.1 Interviews

In the following section, the results of the conducted interviews are compared with other interviews within the group of the same organizational company type. Due to the fact that most of the interview partners asked to handle the information they provided in an anonymous way, they will not be named explicitly during this section. As for the analysis of the interviews, we therefore use the following abbreviations for the interview partners in order to be able to assign the job title to each statement:

- A: Data Governance Coordinator (public organization)
- B: Head of Application Development (public organization)
- C: Team Leader Data Governance & Information Provision (public organization)
- D: Employee in the Strategic IT (large company)
- E: Market Risk Pricing Specialist (large company)
- F: Head of IT and Compliance; Data Privacy Officer (large company)
- G: Chief Data Scientist (SME)
- H: Software Developer (SME)
- I: Chief Executive Officer (SME)

The reason for these relatively diverse roles within the different organizations is to find as much challenges and opportunities in the context of Data Governance as possible. Consequently, it was possible to have opinions from 1) various perspectives within different organizations (e.g. from an employee who is only indirectly affected by the establishment of Data Governance in his department) and 2) thus having a broader view on the current situation in diverse organizations.

At the beginning of this section we give an overview of few general findings from all interviews. The following table states the current knowledge and state of the concept of Data Governance within the interviewed organizations. It can be seen that one third has not heard about the concept of Data Governance before. Another third is already implementing it, and the last third is between these stages.

Table 3: Knowledge level of Data Governance of the nine interview partners.

Knowledge level	**Large companies**	**Public organizations**	**SMEs**	**Sum**
Unknown	1/3	1/3	1/3	3/9
Known	0/3	0/3	1/3	1/9
About to implement	1/3	1/3	0/3	2/9
Implemented & progress	1/3	1/3	1/3	3/9
Fully implemented	0/3	0/3	0/3	0/9

As a next step, the information gained about the current implementation of the ten data management dimensions from the DMBOK is summarized in the subsequent table. We can observe that independently from knowing the term of Data Governance, its related concepts are present in most of the cases.

Table 4: Knowledge and implementation of the eleven DMBOK dimensions in the nine interviews.

DMBOK dimension	**Large companies**	**Public organizations**	**SMEs**	**Sum**
Data Quality	3/3	3/3	3/3	9/9
Metadata	3/3	3/3	3/3	9/9
Documents & Content Management	3/3	3/3	3/3	9/9
Data Architecture	3/3	3/3	3/3	9/9
Data Security	3/3	3/3	3/3	9/9
Data Storage & Operations	2/3	3/3	3/3	8/9
Data Warehousing & Business Intelligence	3/3	3/3	1/3	7/9
Data Integration & Interoperability	2/3	2/3	2/3	6/9
Reference & Master Data	2/3	2/3	1/3	5/9
Data Modeling & Design	1/3	1/3	3/3	5/9
Data Governance	2/3	2/3	1/3	5/9

In a nutshell, although most of the dimensions of data management are already implemented within these organizations, the term Data Governance is not used in all these cases. In fact, in all nine organizations interviewed the formal criteria for implementing a Data Governance strategy were more or less fulfilled. The dimensions which are the least often implemented are "reference and master data", "data integration and interoperability" and "data modeling and design". These three dimensions are mostly based on the other dimensions and can be therefore interpreted as a more advanced level of Data Governance. One reason for the latter being relatively underrepresented is that some interview partners mentioned to not need this dimension since there already exist numerous out-of-the-box solutions for it.

3.1.1 Public organizations

In total, Data Governance in public organizations can be seen as a rather new concept. The need to implement an effective Data Governance program in these cases does not exist for longer than three years. A particular problem indicated by all three interviewed public organizations is that they suffer from historical legacies, which should be managed through the successful implementation of a Data Governance strategy. The organization of A and C already worked on a concrete Data Governance concept, whereas in B's organization it has not been a relevant issue so far.

Both A and C had a different consulting firm at the beginning which has supported them in their earliest phase of implementation. After about six months of consultancy these organizations have created their own specific institution-related framework, which resulted from a detailed analysis of their specific needs and culture in connection with general Data Governance frameworks. With regard to best practices, it can be said that person A did not find any within the German-speaking area, and person C considered only few of them as being suitable.

3.1.1.1 Analysis of the DMBOK

Person C was the only one in all nine cases who has already known this model before. This public organization also claimed to be one of the most-advanced players in the area of Data Governance in Austria due to its long history in dealing with related issues (e.g. effective data management, data quality or metadata).

For person A, the DMBOK has been regarded as a suitable model which depicts important core issues for an effective Data Governance program. Data quality, data security and metadata have been identified as being more relevant than the other technical dimensions and documents and content management. Metadata have been produced for a long time and also published since 2011. With the development of a specialized data model, the implementation of this concept began successfully. At the moment, the organization is working on a suitable business object model, which major parts have already been developed. Nevertheless, it was also stated that it is essential to define the meaning of certain Data Governance dimensions in more detail. For example, metadata can be related to different types of data. Therefore, a useful metadata or master data management is seen as a prerequisite for Data Governance. Moreover, data quality and data should be described in such a way that everyone is familiar with it - i.e. that the data can be found easily, what is summarized under the term of "data descrip-

tion".

In the organization of A, the subject of data warehousing and business intelligence should be rather divided into an organizational level and a technical level. Moreover, data architecture is seen as a cross-section which belongs to all other dimensions. Data storage is rather a technical implementation, data storage and operations have been the longest in existence. Data security is a guiding issue in the implementation of the GDPR and thus already part of its Data Governance program.

In addition, the organization of A is trying to quantify and calculate the monetary value of their data with a so-called "Total Costs of Dataship" approach. The value of the data is determined by the availability and the knowledge throughout the organization. Therefore, the initial step is to define data quality in their specialist data model. Depending on the type of data, this is implemented differently according to specific quality criteria. Poor data quality can for example cause data silos. In the current situation, the data do not fit together and also need to be constantly corrected. In addition, this leads to a lack of further processing in the data warehouse or the business intelligence resulting in incorrect evaluations or plausibility problems (e.g. incorrect key figures).

Person A states that as for the classification of the data there have existed several classes for a long time within the organization, such as confidential, restricted or public data. Related rules and policies exist on a role-related and technical level and have been integrated into their Data Governance approach as well.

Within B's organization, the metadata structure is fixed, but the quality of the input given by the employees is not surveilled or commonly regulated. Although there exist common rules for it, they often differ between the individual departments and are not fully respected. For employees from other departments it can be therefore difficult to work with these data. In addition, no particular meaning is attributed with master data so cross-departmental value of data has not been recognized yet.

3.1.1.2 Implementation and communication

In the organization of A, it was the CIO who made the decision to gradually introduce a Data Governance policy. This is currently being implemented and includes a number of change management processes. It is vital to provide the involved employees with the necessary information, education and training in order to integrate them into the work process. In parallel, the first databases are being adapted by describing reference data stocks and

integrating them into the data warehouse.

For A, the issue of business intelligence is a key factor when dealing with Data Governance. Their implementation phase has started about one year ago. At the present time, several pilot projects are running within the organization with a focus being on the departments that are particularly data-intensive. The necessary time frame for each pilot project is estimated two to three years. Within the next year, the guidelines and training concept should be finalized, and their implementation should then take two years. Success measurement is carried out using a maturity model which has been adopted to their specific needs.

The institution of A prefers a Data Governance model with seven dimensions, which can be mainly seen as a top-down approach and occurs in their corporate strategy. It includes the following dimensions: data excellence organization, business data model, data quality management, reference and master data management, demand management, reporting principles and data risk management. Data excellence is the name of this project and describes the technical implementation for the timely provision of reliable administrative data in the required quality. The data excellence organization itself is the institutionalization of Data Governance. The reporting principles are necessary for efficient evaluations, for example when the data are stored in the data warehouse. Data risk management mainly covers data security, data protection and related topics. The metadata dimension relates to the business data model. It is also planned to build a three-stage maturity model based on this structure.

In the case of C, the concrete designation as "Data Governance" has emerged from a project on how to best prepare data, to create an overall view and to provide information about the entire business process data, data quality and data requirements. Therefore, the board had a clear commitment to the introduction of Data Governance, what is important since it cannot be introduced bottom-up in the opinion of C.

For C, all dimensions of the DMBOK, data quality and metadata in particular, should be integrated into an entire Data Governance model. Their own Data Governance model consists of different frameworks, which - with a half-year support of a consultancy firm - led to their own framework, the so-called Data Governance House. As for the organizational implementation in the case of C, it exists on the highest level of management: directly below the executive board can be found the Data Governance Board (responsible for the outcomes and improvements of Data Governance), which is composed of the individual business unit managers and the CIO. The CDO will be also part of this board. One level below are the Data Governance coordinators who represent the contact persons for organizational issues in their specific

business unit. Within the individual divisions, data owners (data controller for the technical data group) define and take responsibility for data quality within their domain, and data stewards (data set representatives) who implement the requirements of the data owners operationally, can be found. They are further supported by other units of the organization such as controlling. In the last level, operational data managers (regional contact person for the data quality in his specific area) and operational data administrators (responsible for the operational data maintenance) will be available, and data consumers can be found in the individual departments. These people have to be professionals and therefore prove to have experience with their institution. They are further supported by master data manager admins or process managers.

In the case of B, the concept of Data Governance has been completely new. It is not planned by their CIO to implement such a policy.

3.1.1.3 Roles and responsibilities

The person responsible for the whole implementation in the organization of A is the CIO who leads the steering committee for data excellence, the executive decision-making unit of the organization. The subsequent Data Governance coordinator is the leader of the so-called Data Governance advisory board (mainly responsible for the strategic orientation), which is composed of employees from a specific department as well as some of the data stewards of various departments. They participate in the creation of a data quality strategy and development of data quality principles. Moreover, there will be data experts in the individual departments who care about their respective data. They further encourage the data users in understanding and implementing the professional data quality requirements. Moreover, they support the technical design of the implementation of professional data quality tests as well as the release of short-term data corrections, and they will also be the contact person for open government data issues. The data users will be responsible for the enrichment and collection of data in an appropriate quality. Data consumers are internal or external stakeholders who can have access to specific data, depending on their respective needs. Data users are those who use the data, which can also be Data Scientists. To some extent, Data Scientists are available for creating predictions such as population forecasts, and should also deal with predictive analytics and related methods in the next few years.

In this sense, the roles of data stewards, data experts, data users and data consumers have not been implemented so far, but should be represented in

every department with at least one person, e.g. for personnel data, financial data or energy data. The main focus will be put on existing personnel who will then acquire these additional roles for their area. In case of being unable to solve a problem at the lowest level of the Data Governance model, the person concerned asks the level above. A final decision is taken by the steering committee at the latest.

Some other roles already exist, such as the contact person for open data or a Data Protection Officer, who will then get the role as data experts. There has been a Chief Data Officer (CDO) even for seven years, but he is currently dealing with open data topics and data releases. In the future, his tasks will be expanded by the basic tasks normally executed by a CDO.

Beside the roles already mentioned in the section above, a Chief Data Officer at C's organization is currently in the introductory phase and will be on the same level as the CIO and will take over the leadership of the Data Governance organization. Moreover, Data Scientists are rather operational, very focused on their specific area and considered as so-called master data managers in the roll concept. They have a broad systematical and process-related knowledge and can also work with raw information. In addition, they have to keep up a high level of communication amongst each other in order to cover cross-procedural expertise. A total number of about 100 people will be responsible for this Data Governance organization in total. In total, their implementation is nevertheless a challenge because resources are scarce.

In B's case, there exist no Data Governance roles such as a CDO, and there are even no Data Scientists. However, B states that Data Scientists would be able to generate value, but also questions the extent to which it is currently desirable within their organization. In fact, these analyses should necessarily accompany a strategic commitment in order to draw conclusions from the realization.

3.1.1.4 Concepts of data storage and operations

In A's organization, the concept of Enterprise Linked Data is known and is already used and published, but it is not used on a daily basis because the common know-how is still missing. On the other hand, projects in the scope of Big Data and Data Lakes are planned. For example, the concept of a Data Lake is currently being set up on a technical level.

In the case of B, the data of individual managed tools exist in singular solutions and therefore cannot be bundled and evaluated together. Therefore, there are some Enterprise Linked Data and Big Data projects, whereas Data Lakes or Data Spaces are not planned to be implemented. It is the aim to

have data available at different points of the organizations through suitable references, and to link to other areas and institutions. Nevertheless, there is no general strategy for this by now.

For C, these topics belong to the research department. The basic concept of Data Lakes has been present in the organization for many years and is used for example for their measurement data. Another example is the financial department which considers raw data from the Data Lake to find data patterns. Furthermore, C does not see the boundaries between the individual concepts. As for Enterprise Linked Data there is currently a pilot project running which is at its early state not intended to provide data to other departments. It also needs a lot of time to make required software applications available.

3.1.1.5 Data security and data protection

In the case of C, data security and data protection were not the main reasons for the implementation of Data Governance since these issues are currently rather strategic than operational. However, a Chief Information Security Officer is involved in related topics like data transfer or data processing.

On the other hand, it has also been noted by B that due to security concerns, some services are relocated to the cloud. In addition, there is a separate security group which is also responsible for the entire internal area of data protection and data security, as well as for external EU agendas such as the GDPR. At the moment, there is no Data Protection Officer, but his implementation is currently being discussed.

3.1.1.6 General challenges and opportunities

The project in case of C has been running for two years so far. According to C, the size and complexity of the business processes is also a driving factor for the implementation. Their self-created maturity model consist of five stages. Currently, their level of Data Governance is in between the stages two and three. It will still take some time to work, because it means a great deal to bring about these changes. The theoretical concept is already well developed, but the concrete implementation in the individual processes and in the data has not reached an advanced level by now.

Another particular problem mentioned is the loss of information caused by high staff turnover, which should be counteracted by Data Governance. In fact, their knowledge should be extracted to well-defined data structures.

In the opinion of A, the biggest challenge is the speed with which new concepts come into the organization as well as the amount of data increases. Moreover, in order to evaluate these data volumes appropriately, Data Scientists are needed. With regard to the topic of specialist data management, the corresponding tools are currently being implemented for the management of metadata and reference data. In addition, it is a problem that there is too much everyday business to pursue such new strategies with the necessary consistency. Likewise, such organizational changes are usually tedious because people tend to be more afraid of this change and do not see the opportunities. Consequently, it must be explicitly supported by the management which has the ability to make the required resources available and, if necessary, also work with an external consultancy.

3.1.1.7 Strategic importance, goals and overall consideration

In all three cases, it has been agreed upon seeing Data Governance as a hype-issue, although the basic ideas behind this concept have already been in place for a couple of years. In fact, it is stated that Data Governance has been a topic for a long time. It seems to bring a change theme into the organization. This means that it reaches both technology and the organization.

In the opinion of A, one major aspect of Data Governance is that it is used from an organizational point of view. It is important that Data Governance is implemented with a top-down approach and that there is someone who has a clear responsibility for it. It should be supported by the management, since this is the only approach which ensures a continuous and serious implementation. However, a simultaneous bottom-up approach is needed to reach the people within the departments and to inform and motivate them for further training courses. It also requires a common understanding and interpretation of its key concepts such that those who are addressed can work with it. Therefore, the employees are informed about the topic through various internal media like small reports in order to raise the overall awareness right from the initial phase.

For A, the overall goal would be that all data in the data model are described and stored in a particular repository, that there are no more data silos, that clean reference and master data are managed, and that evaluations can be made immediately from these data. Person further A supposes that the concept of Data Governance is in general rather underestimated, whereas it should already be implemented on a more advanced maturity level in the banking sector.

The focus on the business aspect is essential for C. Overall, it is seen

as an organizational and cultural project, not only an IT or a pilot project. One goal is to integrate data maintenance into the business processes. This is based on their self-developed framework, which includes the four major areas: data, organization, processes and systems. For C, Data Governance is further considered to be a continuous process. The overarching goal is to create a stable organization and to manifest Data Governance with the business processes. A time frame of two years is set, with the main difficulties in the creation of a suitable business data model and in the processes to ensure process support. Other aspects that should be taken into account in the future are information transfer or open data topics.

A crucial qualitative requirement for C is that information and data within the business processes are available at the right time for correct application in the right format and quality. Furthermore, these aspects should be ensured over the entire business process ("fitness for purpose"). The major problems are related to the initial phase and mainly consider professional resources. Data quality is determined by the data quality model with 15 parameters, similar to the one provided by Wang and Strong (1996) [77].

B states that the value of their data results from their quality, which depends on the correctness confirmed by a special signature. Poor data quality results in a bad reputation with external stakeholders and could therefore be better managed or even avoided with a Data Governance approach.

Person B is of the opinion that a fundamental review of the organization, the tasks and the data could lead to the introduction of a Data Governance policy with all its related benefits. However, without a convincing cost-benefit weighing, it cannot not be possible to raise awareness of the general management or other employees about the new roles of a Data Governance policy. With an increase in the amount of data, this pressure could result in a Data Governance program.

3.1.2 Large companies

As for the current situation of Data Governance within large companies in Austria, three different options derived from three cases: piloting Data Governance, following a different approach but fulfilling the basic ideas of a Data Governance program, and already having it introduced within the company on a more advanced level.

3.1.2.1 Analysis of the DMBOK

In the opinion of E, one aspect of the DMBOK which should be taken into account is the origin of data. This could occur in the field of data quality

or in a separate area depending on which data are considered. However, the dimension of data processing itself in terms of what to do with the collected data is not represented and should be added to the model accordingly.

In the opinion of F, a major critique on the DMBOK is that the related business activities are rather underrepresented. In fact, it seems to be more a technical model, which ends where an IT system has been technically made available. According to the opinion of F, the model should be adapted to the business-related relevance and as a result to focus more on the business aspect.

In the company of F, the main concept of Data Governance has not been well-known so far. Nevertheless, there exists a solution to this problem which follows a different approach. As an alternative to a Data Governance strategy, there is a so-called business partner organization in place which is part of the entire IT. Their tasks mainly involve mediating between the business and the IT division.

The definition of data ownership within F's company is very important in order to know who is responsible for the technological infrastructure throughout the whole company. According to F, Data Governance in large companies does not depend on a single organizational unit or a single person, but should be shared between three elements, namely: IT provision, IT management and business.

In the case of person D, metadata and data security are essential aspects of their Data Governance program. Person D is working in his company on a pilot project in Data Governance. This project is especially important since it is about customer data and therefore also related to keeping a good image.

3.1.2.2 Implementation and communication

A clearly defined field within D's company is regarding a certain kind of customer data of another company within the group, which value and business relevance has been recognized. Since the project is conducted from the view of the strategic IT, no operational tasks are considered. In particular, it is about the creation of the necessary regulations for how to deal with the data in the company as a whole, i.e. what roles and related tasks, obligations, competencies and processes have to be defined and how to deal with data in principle. Approximately 20 employees have been working with about ten percent of their working time on this project over the last three months. Moreover, an external consultancy firm is currently supporting the achievement of the defined goals, but there is no specific model to be applied by now. The focus of this pilot project lies in the area of data quality and in

the context of how we link data over company boundaries within the group.

So far, in the case of person D, they have been working on the theoretical concepts of Data Governance, what will be further continued over the coming six months. This project is then intended to serve as basis approach of the issue for all essential data stocks. At present, there is still a strong system-driven perspective of the IT or the applications on the data. There is no view of the data per se, and therefore the attempt to create the awareness that data are important and that they need to be managed accordingly exists, independent from the application they relate to.

In the case of the company of E, the concept of Data Governance is already implemented. There are many areas within the company that work directly with this topic and also manage other types of data. With the purpose of Data Governance, E also provides data quality assurance, especially when there are many different data sources, for example to prevent duplicates and to have a central data source all departments can access. The employees are in most cases the already experienced employees who are familiar with the systems and their interplay within the company. In general, the cross-departmental communication works well since the definitions which are worked with are identical. Nevertheless, there are still topics where the understanding of the data is different. In such cases, the data stewards are needed to provide a mutual understanding on both sides and to solve the specific problem.

The company of F mentioned to have a specific unit within their IT department, which is in the position to communicate with the IT and the other (business) parts of the company concerning larger IT-related issues. They represent the idea of a Data Governance program, which also contains all ten dimensions of the DMBOK, which are implemented and revised, and can be assigned to one of the three areas mentioned above. However, this company does not use the term Data Governance in this context. It also lacks the related separation of roles, for example it has not defined positions like data stewards or data experts. This is due to their special organizational architecture since they operate as a huge global corporate and therefore argue that Data Governance in this sense is not deemed necessary in this context due to the complexity of the whole organization and its processes. Therefore, the communication between the different departments costs a lot of time since there exists no common tool which supports their data usage. However, it should be remarked that one reason for it are legal constraints which exist since the legal system does not consider companies in different countries as a corporate but separately.

A considerable number of Data Scientists is working in a global team for a long time at the company of F. There is also a CIO in place, whereas there

is no need for a CDO according to F because in the end, this person has to do what the business departments are already doing. There is not yet a dedicated Data Protection Officer established, but its introduction is being discussed, although not in turn of the implementation of a Data Governance strategy but resulting from the initiative of the law department.

For F, there is in general a possibility of introducing a Data Governance Officer who centrally unites the three aspects of IT provision, IT management and business. In a large company, however, this would not be implemented in the representation of one person because each department within this company would have its own data-related topics. Consequently, F states that if there would be one responsible person in each department it would be too complicated to manage due to the size and complexity of the whole company. In a medium-sized company, it might make sense to name someone who is responsible for all the content as well as for IT, and also for security issues.

Communication between the individual departments within the company of F is divided into a three-tiered system, which consist of an executive management, key users and end users. Key users are people who show a very strong affinity for IT topics. Such people have always been important, because they understand both the IT and the business operations. The end users are responsible for support functions, help desk, masterclasses or trainings. In order to communicate IT topics and other topics to the business departments, one needs a concept for each of these three levels; each of them has to be adapted to the different needs of these departments.

3.1.2.3 Roles and responsibilities

In D's company, there will be data owners, data stewards and other necessary roles implemented, but more likely in the form of a personal union due to the scope of the very small area covered by the pilot project. It is rather unsure whether their Data Governance approach will also work group-wide and where it is pursued, even if there would be a central body governing Data Governance across the entire group or in each sub-company.

The creation of new jobs due to Data Governance is something that is not realistic by now in the opinion of D since it is generally tried to use existing employees as they are experiences personnel who deal with these topics and data. These existing employees already bear responsibilities in their decision domains, not as data owners or data stewards, but rather as process owners who are responsible for the processes of these data, and will then also take over these roles for Data Governance. Whereas a CIO exists, a CDO has not

yet been introduced, and Data Scientists have been in use for several months in different departments.

Person E is of the opinion that their Data Governance policy is already relatively advanced in their company. Specifically, this is supported by data stewards who are already present in the individual departments and are building the bridge to the IT department in case of major IT-related problems. In the department of E, one data steward comes on nine employees. They are also active in the field of data management systems and are therefore responsible for a broad range of tasks. However, these employees have already been working in their department and have gained this new role due to their experience. Next to the CIO, a CDO has already been implemented as an existing executive role.

Within F's company, responsibilities are clearly separated between the departments without any relation to a Data Governance program. There are for example people from the business operations, a compliance department that looks at data protection or a business partner organization in the IT which looks at the security and on other related topics and bridges the gap to the business activities. All these roles exist, but they are not combined or strongly related to each other as it is likely to be in a Data Governance organization.

3.1.2.4 Concepts of data storage and operations

As for person D, some Big Data initiatives already exist within the company, for example for dealing with a huge amount of machine data, produced by power stations, vehicles, energy management systems and customer data. Moreover, the concept of a Data Lake is planned to be implemented, whereas Data Spaces and Enterprise Linked Data are not used or known by now.

The data management system in the organization of E is similar to the principle of Linked Data. There is a source for a certain kind of data, and any user of the data in the company can have access to this specific data to use it in order to avoid redundancies and duplications. Big Data projects are also existent, especially in the area of data mining in very large database systems.

3.1.2.5 Data security and data protection

In D's company, Data Protection Officers and a Chief Security Officer with regards to data security and data protection already exist as an internal service provider with around twenty people. Furthermore, the GDPR is

influenced by the before mentioned subproject since it is part in the overall project which is handled by another department. An important argument for the introduction of Data Governance is to be able to better deal with future legal data-related challenges. Data Governance is not planned to be used for the implementation of the GDPR, although it is seen as a good reason for supporting such projects in the future. If the project succeeds, it would be expanded at first for customer and personal data throughout the entire group.

In the opinion of F, data protection (i.e. all the agendas of the data protection authority, data processing registers) has nothing to do with Data Governance within their company. Data Governance rather includes concepts such as information security, technological security and access security.

3.1.2.6 General challenges and opportunities

The costs of the whole project in D's company have not been estimated so far. For D, it is still difficult to define how an objective before-and-after analysis might look like. It will, however, be relevant to find a suitable way until the end of the project in order to convince decision-makers about its importance and continuation. The goal is to have a set of rules for all the data areas that are essential to the project and to know how to handle these data. At the moment, the project is executed by the strategic IT and is therefore being seen as an IT program. In the course of time, however, it has to become a business topic because these topics cannot be treated within the IT alone.

Within the company of person D, the implementation of an enterprise architecture management tool is being tested in parallel, which will allow the mapping of data structures and will be adapted by introducing concrete Data Governance rules. Data quality is very important or at least a necessary prerequisite to be able to combine the data in a meaningful way and to prevent data redundancies in these systems. Another advantage for D is to have clean data stocks.

In particular, the documentation of data-related individual processes and process flows is important and already exists for the major part in E's company. This prevents the use of methods that are not officially agreed within the company. For this purpose, working instructions are used, which, for example, determine whether documents can be updated or not. Furthermore, there are already concrete changes in the area of data architecture, i.e. how the data are mapped and stored in the specific systems. This is demonstrated by defined nomenclatures. For example, different market data points

are divided into objects, which are then summarized into groups. There are also strictly given classifications, and when there is a need for a new one, there is also a specific process for this development available. Moreover, a dual-control principle exists for the area of data security and data quality.

An important prerequisite for a Data Governance is also a realistic time management with appropriate buffers. For E, the quality of the data is mainly defined by its correctness. Incorrect and incomplete data can lead, for example, to the calculation of incorrect key figures, which subsequently harm the business with other companies and their own reputation.

The major problems in Data Governance for F would be the data architecture, someone who ensures the quality and compliance and the cohesion of the whole portfolio. This is usually no longer exclusively the IT alone. For their company, F states that the pure rendering of technology loses importance because most technical problems are solved. What is not solved is how this can be used in an optimal way. For this, the approach of Data Governance can turnout to be appropriate. However, a few prerequisites must be met which are not applicable to their company. The problems that F is able to solve are, on the one hand, the maturity of the IT products, networking, Big Data, cloud, and the size and complexity of the company. Besides, national borders will play less and less a role for entrepreneurial activities.

It is stated by F that it can be useful to introduce and institutionalize Data Governance and for the distribution of new responsibilities. F is also of the opinion that a potential Data Governance department would be very similar to their existing business partner organization, which in principle could handle much of that. One thing that could be additionally considered is data protection and that there could be individual responsibilities for Data Governance - a contact person - for each organizational unit within these departments.

Another challenge in the eyes of F is that Data Governance could lead to a strong bureaucratization because there are already roles for data quality responsibility. A potential role or responsibility for Data Governance could make the current situation even more complicated since the employees are already relatively busy and an additional role might not have the desired positive impact on the work processes.

3.1.2.7 Strategic importance, goals and overall consideration

According to D, the importance of Data Governance is certainly underestimated, and some companies are only at the beginning of developing a systematic Data Governance approach. The management's awareness, e.g.

that it can support issues such as the GDPR might still be underdeveloped yet.

Data Governance is intended to provide clarity to the company of D in projects about where the data comes from, how good their quality is, and whether or how they can be used. At the moment, it is rather difficult and needs to be redefined for each project. Therefore, these structures are to be introduced, such that these processes can be implemented faster and more efficiently.

IT and business operations are both considered as being relevant in the case of E as the IT is needed to provide the infrastructure being also responsible for customizing solutions. The business activities are mainly designed to make changes in the IT and to specify general adjustments.

According to E, the difference between large companies and SMEs is that they can hire new staff for these responsibilities, and that Data Governance can be lived more explicitly than implicitly within larger companies than in the SME sector. E is the only one of the nine interviewees to believe that Data Governance has already gained consciousness and that the importance is recognized within the different industries. This assumption may be influenced by the content situation within his own company. On the other side, E has no information on the implementation of the GDPR.

Nevertheless, the company of F always needs technical experts from the IT perspective as well as the professional experts from the business operations. F also notes that Data Governance is not institutionalized, but the individual tasks are present in their departments, partly in the IT, business or compliance department. All that exists, not at the same level but rather with different focuses, and sometimes only implicitly. On the other hand, F states that Data Governance would enhance data quality and compliance, but there is the danger of the complexity of these processes, which are present in a large company. For a medium-sized company it might be a good idea.

3.1.3 SMEs

As for the SME sector, three small-sized enterprises with a technological background were selected for this thesis. Their number of employees ranges from eight to twenty. One aspect they have in common is that although their roles are clearly divided, the employees support each other during various tasks not defined in their specific role. On the basis of the three cases, it can be stated that the role distribution and the area of responsibility of individual employees becomes more concrete the larger the enterprise is. In the relatively small enterprises of G and H, which have about ten employees, these responsibilities are clearly defined, but the boundaries are fluid. In

contrast, case I already falls into the category of medium-sized enterprises, whereby Data Governance can be integrated into the firm in a more formal and concrete framework.

Regarding the SME sector, person I is very familiar with the concept of Data Governance, and has already done many projects with customers on this subject. The other two persons G and H have not heard of the term yet or used it within their enterprise explicitly.

3.1.3.1 Analysis of the DMBOK

Both G and H have been comfortable with the individual dimensions of the DMBOK and transferred them to their firm. From person G's point of view, both privacy and the management of data are an important aspect. In addition, it would be necessary to add an analytical aspect to the DMBOK model. This includes, for example, the application of data science or machine learning, with which new and valuable insights can be drawn from the edited and qualitative data. For H, the topic of speed should have a larger role in the model. This depends on several factors such as the metadata and the data architecture, but its importance is very relevant to the business.

For company I, semantics are particularly important, which could also be cited as a separate dimension in the DMBOK. A corresponding semantic layer in the field of data architecture could be supplemented with the subdivisions taxonomies, ontology, mapping and querying. According to I, business intelligence is possible at a higher level with the aid of Data Governance because a suitable data architecture helps to integrate the data. By linking to the metadata, an interpretable semantic model can emerge. Therefore, person I sees it as a layer model rather than a separate differentiation according to their importance. Nevertheless, individual layers are building upon one another and unfold their potential only in conjunction with the other dimensions.

Poor data quality can be expressed in many ways, according to G, especially in missing or incorrect data entries. One negative consequence are incorrect in-house analyzes, which would have a negative impact on their main product and business. In H's enterprise in which the business model is also determined by a high data quality, incorrect data would have a negative effect on their reputation. This once again underlines the high importance of appropriate data management in connection with the strategic orientation of the organization.

The company of H also considers the sub-range of data quality. In order to measure data quality of individual data objects, a metric has been defined,

and any modification of the software must be tested before it is released. The main success parameter is when the measurable data quality after a change of the software is better than before, and at the same time more records can be processed.

3.1.3.2 Implementation and communication

The introduction in the organization of person I has started when they were dealing with projects of large customers for the first time. For example, their customers use their product to improve one specific aspect of their Data Governance program through professional metadata management.

With regard to the cross-departmental communication and implementation of implicit or explicit Data Governance policies, it can be said that in the opinion of G and H, Data Governance works particularly well in small companies and start-ups. H also takes the view that companies in which communication between individual departments is not possible because of strongly separated areas, Data Governance does not work well.

Overall, person I states that most companies in the German-speaking countries still have a too strong separation of IT and business operations. The data stored in the IT departments of the organizations are only poorly involved in the business processes. As a result, many organizations face great difficulties in competing in the competitive market environment.

3.1.3.3 Roles and responsibilities

The role of Data Scientists for the preparation and analysis of the data is present in these three cases. However, as person I emphasizes, this seems to be the exception and most companies, if they are not specialized in this field like G and H, are not yet represented.

With regard to the roll allocation, person I notes that the company does not have an explicit distribution such as the position of a Data Security Officer. Rather, this is implemented by an embedding in other roles. This can be also seen as a difference to larger companies. Likewise, in all three cases, the role of the CIO exists, but in the small enterprises, this role also coincides with others. A CDO does not exist in any of the SME cases, whereby person I mentioned that he has learned from his customers that a CDO is strongly represented in data-driven branches (for example in the media and publishing sector, but also in the pharming sector), because in these industries the efficient and precise management of vast amount of data is required through the gradual emergence of personalized products. For

this, the concept of Data Governance is very suitable and present within these organizations. In other areas, such as construction engineering, this view of the data has rather been neglected so far.

A Data Protection Officer already exists within company G in an even more advanced form as specified by the European Union. For H, data security and data protection are less important since the majority of the data they use are freely available to the public. In addition, however, the sensitive customer data has to be protected by appropriate measures. The company of H is much more precise than what the GDPR prescribes. It is therefore also strongly depending on the business purpose of the company to which extent the data security aspect is relevant. The company of I is currently dealing with this directive and says that it could be the case that there will be further personnel adjustments.

3.1.3.4 Concepts of data storage and operations

In the case of company G, the concepts of Data Lakes, Data Spaces and Enterprise Linked Data are not used, but rather dealt with them mathematically, especially with Big Data. As for H, only Big Data is known and applied within the enterprise since Big Data Governance can support a successful and efficient use of several million records and the related metadata.

3.1.3.5 Data security and data protection

As for data security, care is taken whether there are unintended ways to get to the data. If such a tested security attack fails, the security tools are appropriate. Data security with the essential sub-items privacy and data protection are closely related to the legal framework, especially with the GDPR. Certifications and ISO standards are an important additional qualification for both G and I companies. Overall, the implementation of the GDPR is stronger combined with an implicit Data Governance policy than in the area of large companies and public organizations.

Due to outsourcing of the webserver, one does not have to worry about the security of the servers in the case of H and therefore also have less administrative effort in their own company. Thus, the effort associated with a comprehensive Data Governance policy can be reduced by means of a corresponding outsourcing of services. The company of I is concerned with these concepts in customer projects, but does not need them for its own product development or administrative tasks.

3.1.3.6 General challenges and opportunities

With Data Governance, person I believes that one big opportunity is that companies can continue to grow. In doing so, data is deliberately perceived as a valuable raw material or as a value-adding element and incorporated into the business processes.

G sees the concept of Data Governance as a possibility to make IT terms concealed behind this topic more tangible to people, as it is also the case, for example, with Big Data. For smaller companies, such an explicit implementation is not necessary at the beginning, but the individual dimensions of the Data Governance framework are. Only with increasing size and complexity of the enterprise do all three persons see the possibility to distribute roles of Data Governance more clearly. Person G sees this limit with approximately 100 employees.

3.1.3.7 Strategic importance, goals and overall consideration

For the measurement of the success of a Data Governance program, person G proposes a before-after analysis in the area of data security. By preventing the theft of sensitive information and comparing it with other companies where breaches have happened, one could estimate the cost savings and value in general. This could also be applied to the field of data quality. If existing data are improved, a value is generated at the end of the chain, which can be quantified according to the use of the data. On the whole, however, person I concedes that concrete measurements are also a major challenge in his company.

Moreover, the unifying function of Data Governance is emphasized once again in all three cases. According to G, Data Governance is a hybrid of IT and business, with the business playing a central role. Data Governance and their terms are referred to as "a business taxonomy for IT topics". Above all, however, the business aspect is particularly important because it affects the company as a whole. However, large parts of the previously defined requirements are implemented by the IT. Person H sees more the business aspect in the foreground, whereas person I does not recommend to make such a strict separation.

All three are of the opinion that Data Governance is becoming a business-relevant hype topic, which potential has not been sufficiently recognized by now. Person I states that "today, in 2017, we are writing the time where every company should have recognized that money and added value is to be fetched in this area, even in the dusty governmental sector". In addition,

H is of the opinion, that the importance of Data Governance is generally underestimated, especially in the legal and security sector. Person I also believes that Data Governance is generally not given enough importance.

The overriding goal of Data Governance for G is that all organization-relevant Data Governance processes that occur in the area of data and the processes that pass through data should be ultimately covered by clear responsibilities. The long-term goal for H would be to make the resulting benefits measurable. For this, it is decisive to be able to measure how well the data are related to each individual dimension of the DMBOK. If this can be measured, concrete improvements are possible. For I, the goal is also to obtain certifications from authorities about the status of their achieved Data Governance level in order to make it visible to external stakeholders.

Central to person I is also the commitment given by the management of the company. If Data Governance were only operated by a department, it would not work. In terms of implementation, person I believes that an immediate top-down approach is not realistic. Instead of a single pilot project, for example, several projects in the areas of data architecture and reference and master data should be executed. Started in these areas, experience values would arise, and with the resulting knowledge and a suitable strategy, this incremental process would be continued. The duration of a comprehensive introduction of Data Governance can take up to ten years, depending on the intensity and position of data in the respective company.

3.2 Online questionnaire

In this section, we consider some findings from the online survey. It is important to state that the information concerning the public organizations are not valid enough since in total, only four of them participated in the survey. Nevertheless, the outcome for this type of organization is considered. In total, 44 organizations have taken part in this survey, with 33 from Germany (78.6%), 7 from Austria (16.7%) and 4 from other countries or not specified. They are categorized by the following industries:

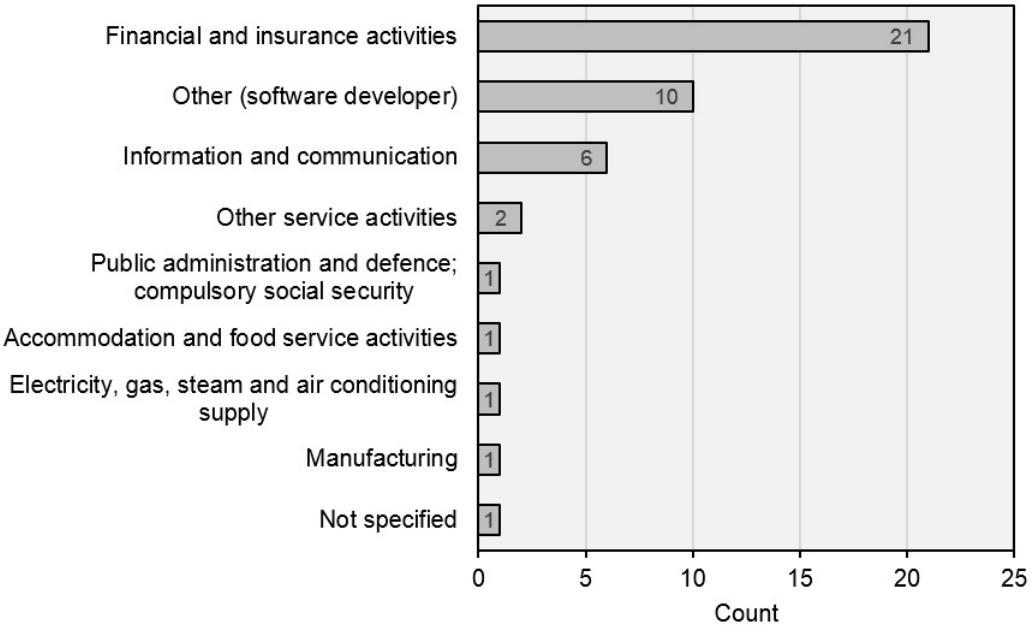


Figure 4: Distribution of the industries among the participants of the online survey (source: Statistics Austria) [5].

It can be seen that most of the organizations belong to the financial and insurance sector or are software developers.

3.2.1 General results

Overall, 31 organizations (70.45%) said that they have already had experience with Data Governance. As for the definition for Data Governance provided for this thesis and its potential modification, the following comments have been made by participants of the survey:

- "The aim should also be to comply with relevant data protection regulation."

- "I would add: with the aim to ensure privacy."
- "Data should not only used to its maximum extent. Data protection should be a main goal as well."
- "Ok. I would add: Data Governance is a top down topic broken down from top management to the operational units."
- "Data protection is missing."
- "I am unhappy with the definition of the aim. From my perspective it has to be enhanced by compliance issues. There are a lot of data the organization uses, but does not own, e.g. personal data."
- "I would add Data Quality to the definition as this is the key aspect from my point of view."
- "I'm missing the security issue - data governance in my opinion also contains the use of data according to national/international rules concerning data security."
- "What about Data Protection and Security, or is it meant as a part of organizational rules and responsibilities? In my focus it is GDPR."
- "Data Protection not mentioned at all."
- "I miss the security aspect."

Consequently, it can be seen that for most of the participants, the aspect of data security (i.e. privacy and data protection), especially in relation with the GDPR, are an essential part of or for implementing their Data Governance program. In addition, data quality and compliance issues should be added to the definition as well. This view can be explained by belonging to the financial and insurance industry, where regulations are strict and penalties are high in case of a wrong behavior.

3.2.2 Analysis of the DMBOK

Furthermore, the vast majority of the participants (86.4%) has not heard of the DMBOK model before. However, organizations are familiar with many of the single dimensions. The following graphic displays the distribution of the ten dimensions within the organizations observed:

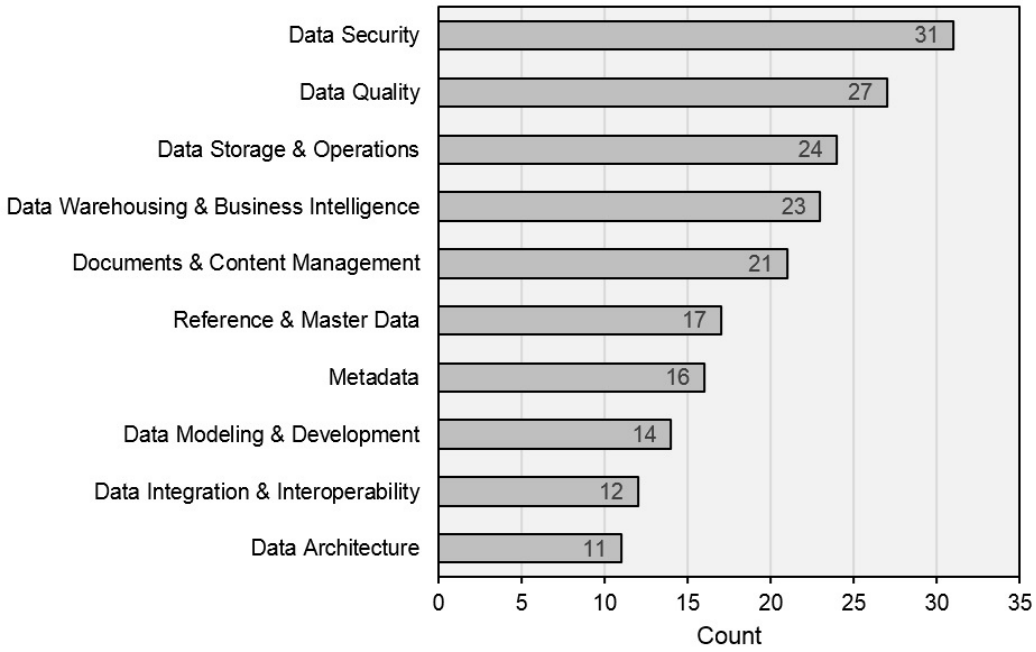


Figure 5: Occurrence of the different DMBOK dimensions in the online survey.

Based on this part of the survey, data security and data quality are the two dimensions which are pursued in most of the organizations. This complies with the results from the interview analysis. In case of data warehousing, data storage and operations and document and content management, both surveys deliver similar results. However, metadata has been mentioned relatively more often in the interviews as in the online survey. This may be due to the fact that the business of the interview partners is more concerned with having a metadata environment and therefore also have started earlier than other industries to use these data structures within their organization.

As for the implementation of a Data Governance policy, the survey has shown that it has existed within 50.0% of the organizations for at least five years, whereas 26.2% have not implemented it so far. Moreover, about two third (63.3%) think that it would take up to five years to implement Data Governance within their organization. On the other side, 31.8% are not knowing how long this implementation could take in total. Therefore, the time factor seems to play an essential part in the decision of whether or not having Data Governance, which also comes along with a large resource input.

As for the question whether the participant's company uses a particular Data Governance framework, the following answers were provided:

- "Part of data secrecy and protection of CID data."
- "X-Gen addresses the challenge of business and technical complexity."
- "SFG Data Model, Secure Data Protection, Big Data Approaches, ..."
- "Data Governance Framework for banks according to BCBS 239."
- "The Data Quality Framework describes the framework of the DQM. It is a tool of Data Governance Management."
- "Policies and rules around the operation and use of our internal data ware house."
- "Contains rules and aspects for data quality of financial data from group companies."

Consequently, it can be seen that a number of the organizations uses related frameworks for single dimensions of the DMBOK, especially for the dimensions of data security, data quality, data storage and operations, and data warehousing and business intelligence. These results can be explained once more with the strict regulatory environment of the organizations observed.

3.2.3 Roles and responsibilities

In this context, the most common roles that are related to Data Governance inside the organizations are in descending order: data user, data owner, Chief Information Officer, Data Protection Officer and data expert. This result is similar to the findings from the literature review. Data Scientists were mentioned to exist in almost ten percent of the organizations.

This means that especially the role of having clear responsibilities for clearly separated data areas in the operative tasks for managing data are very present, whereas their overall managing function is mainly part of the activities of the CIO or a Data Governance committee where the CIO can be part of it. Chief Data Officers as a new role for especially data-related issues on an executive level is rather underrepresented at the moment.

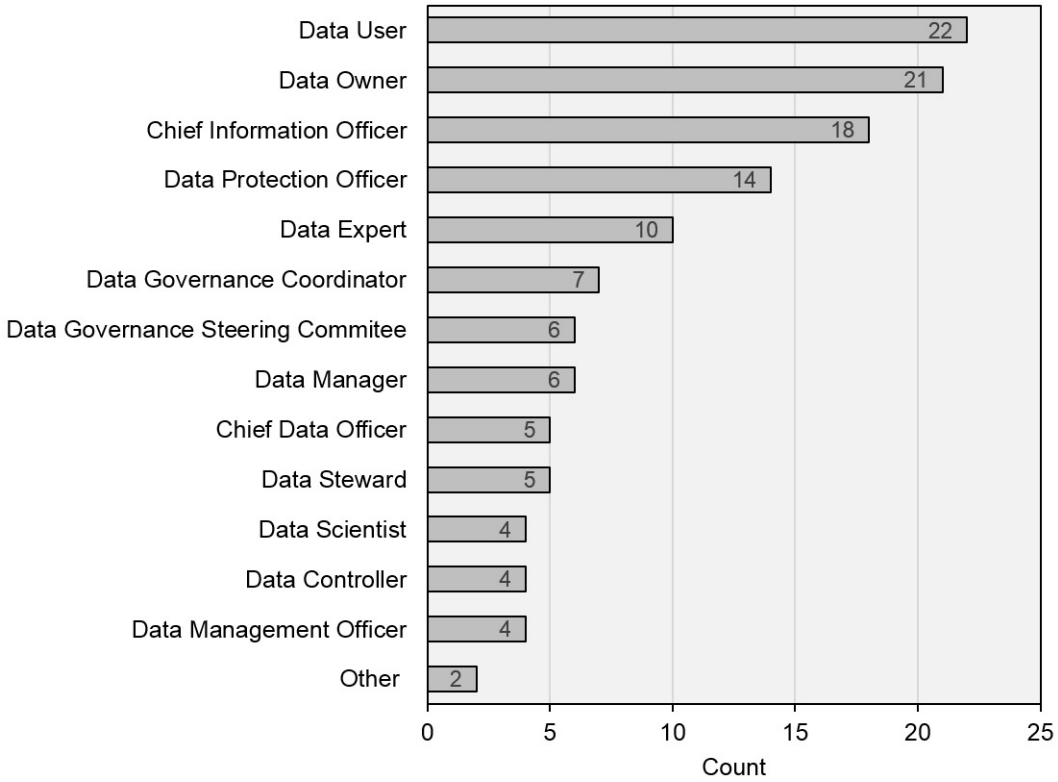


Figure 6: The most common roles within a Data Governance organization of the online survey.

Moreover, 55.8% stated that they would only train existing staff, whereas 20.9% would hire additional staff for the establishment of Data Governance within their organization. These results are similar to the ones from the interview analysis since existing staff already has valuable experience and knowledge within their working areas and should therefore be equipped with these new responsibilities.

In addition, 77.3% mentioned that the meaning and value of data is very important (8 points or higher on a 10-point Likert scale). Therefore, it can be seen that in most of the organizations which are concerned with Data Governance exists an understanding that data is an essential value creating asset which has to be governed appropriately. This view is also consistent with the one from the interview partners showing that Data Governance organizations have or are about to develop a data-driven business.

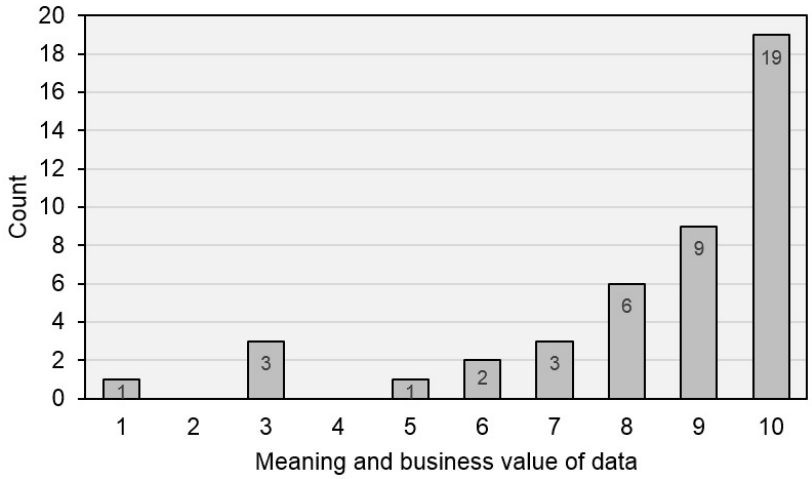


Figure 7: The overall estimation of the importance and value of data within the organizations observed.

3.2.4 Concepts of data storage and operations

This section provides an overview of the results for the data storage concepts in relation to Data Governance. In the survey, the participants stated to have the following concepts within their organizations or are about to implement them:

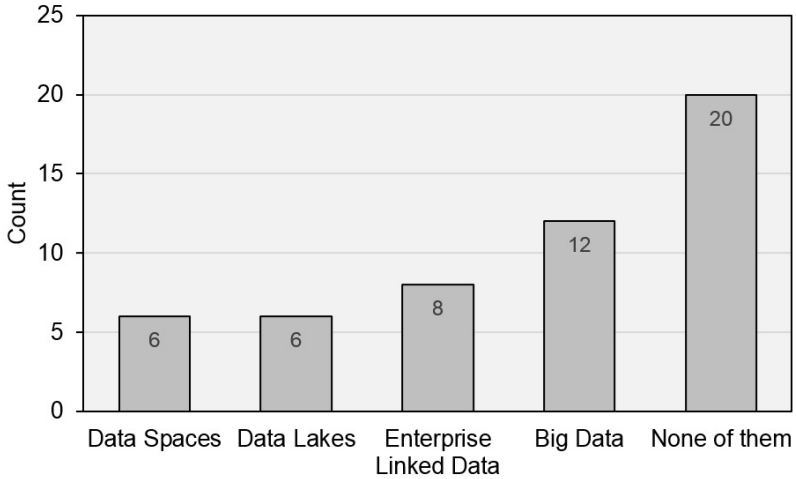


Figure 8: Concepts of data storage and their existence within the organizations.

Another questions has focused on how relevant these concepts might be for their organization.

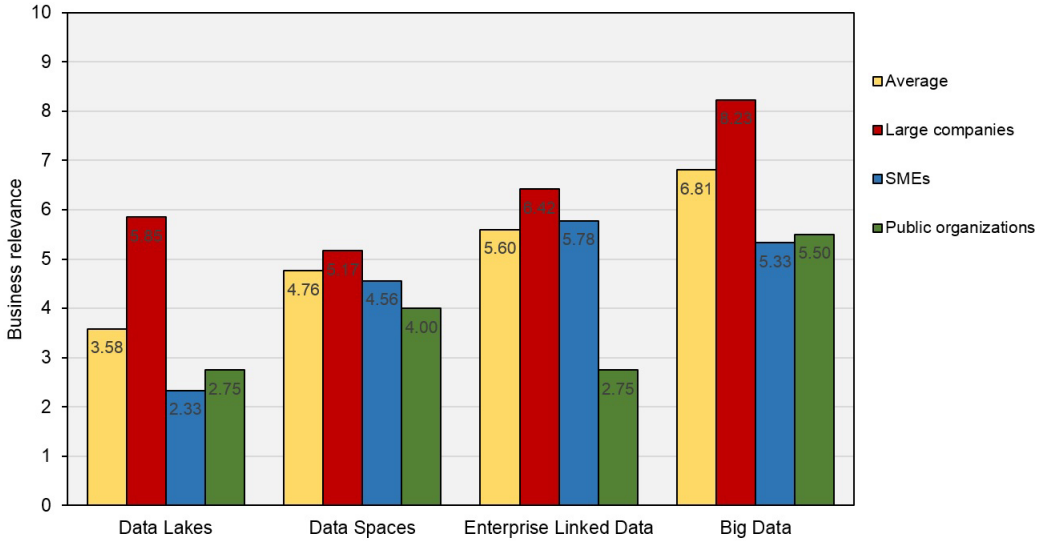


Figure 9: Concepts of data storage and their business relevance, separated by organizational types.

The graphic above depicts the distribution of current projects concerning different concepts of data storage and operations within the organization observed. The results show that large companies are regarding Big Data and Enterprise Linked Data initiatives as the more important concepts, what is also true for the average of the three organizational types concerned. In case of the SMEs, Enterprise Linked Data is seen as more important than Big Data, whereas in all other organizations, Big Data projects are considered as being more relevant. Data Lakes only seem to have importance within the large companies, and the relevance of Data Spaces is equally distributed between the organizational types.

3.2.5 Challenges and opportunities

In total, 27 challenges and 27 opportunities have been observed in the online survey. The full table concerning the average values of the 10-point Likert scale separated by each company type and in total can be found in the appendix assigned to the table with all identified challenges and opportunities from the interviews. Overall, the following aspects have been identified as being the most challenging when implementing a Data Governance program:

- cost intensity,

- definition of all related rules, processes and policies,
- time needed for the implementation,
- lack of best practices and
- effort necessary to ensure data quality.

The overall opportunities that can result from a correctly implemented Data Governance program are to

- have ensured data security;
- meet legal data-related requirements;
- make better decisions based on improved data quality;
- treat data as a valuable business asset;
- have the right data for the right processes;
- ensure regulatory compliance; and
- have consistent data quality policies.

These results show that an effective Data Governance program needs numerous organizational resources, what can be difficult to provide. On the other hand, the desired benefits of data security, legal compliance and data quality as the main factors to ensure business activity within the organizations observed are likely to be promoted by Data Governance.

Therefore, bridging the gap between business and the IT as an important promise of Data Governance is not considered as being one of the largest benefits. This may be due to the fact that this results implicitly from the other opportunities. Another potential reason can be that this collaboration already exists to a relatively large extend within most of the organizations observed.

The following graphic regards the supportive role of Data Governance in terms of how it could contribute to an effective IT strategy, business strategy or to legal compliance such as in case of the GDPR.

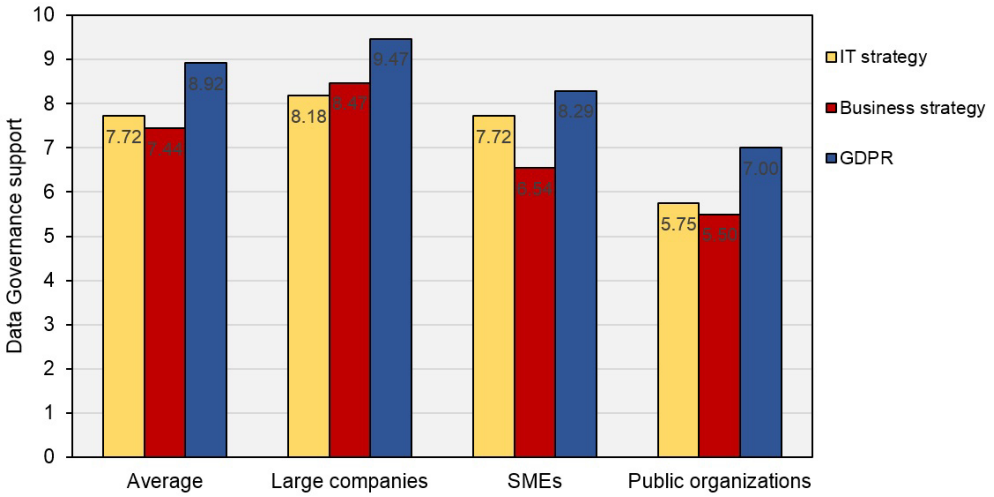


Figure 10: The supportive function of Data Governance.

What can be deducted from these dependencies is that Data Governance can be effectively used to support the organizations observed with the implementation of the GDPR. Furthermore, the IT strategy can benefit a bit more than the business strategy in case of SMEs and public organizations. Only for the large companies, the business strategy could be slightly more supported than the IT strategy. Nevertheless, these two perspectives seem to be rather equally supported by a Data Governance policy, which is also congruent to the findings from the interviews and the recommendations in the existing literature.

In addition, the participants of the survey also had to assess their current Data Governance maturity level on a 10-point Likert scale, with 10 having reached the optimal potential. The results show that large companies have already realized about the half of their possible potential, whereas the SMEs see themselves about two points below the latter, but also two points before the public organizations.

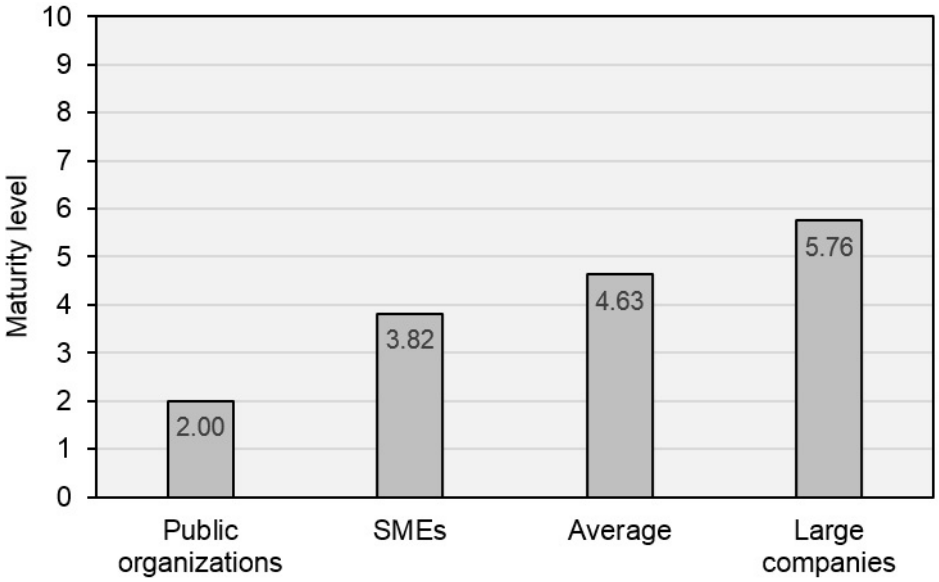


Figure 11: Estimation of the Data Governance maturity level in the online survey.

Last but not least based on the survey, for almost all of the participants, the implementation of Data Governance is not limited to a specific number of employees. This is also consistent with the view in the interviewed organizations: Data Governance should be possible to be introduced in any organization, however, adopted to the different levels of complexity.

3.2.5.1 Large companies

As for the large companies in particular, 85.71% stated that they already have experience with Data Governance; and 38.09% of them even have an own Data Governance department. All of them work across different departments in order to implement Data Governance. Consequently, the value of data (9.24/10) and security and data protection (9.48/10) are considered as very important for their business, with 95.24% stating that Data Governance can support the appropriate implementation of the GDPR.

The most challenging activities for the large companies in the context of Data Governance are:

- cost intensity,
- definition of all related rules, processes and policies,
- prospective legal obligations,

- seeing data as a value-creating element,
- time needed for the implementation,
- complexity of a Data Governance framework and
- effort necessary to ensure data quality.

This order is rather equivalent to the general view for all organizational types, but does not see best practices as a main challenge for implementing Data Governance. Further new challenges mentioned by the participants are to have access and ownership policies and to increase the acceptance of customers for the data use.

On the other side, the most beneficial activities in the context of Data Governance are to

- have better decisions based on better data quality;
- ensure data security;
- meet legal data-related requirements;
- have the right data for the right processes;
- treat data as a valuable business asset; and
- minimize data redundancies.

Also for the opportunities, the view of the large companies is rather equal to the general perspective, but it does not mention to have the right data for the right processes. An explanation could be that large companies tend to be relatively advanced in ensuring data integration and interoperability throughout their organization.

Further opportunities mentioned by the participants are to create new business models (Keyword: "LegalTech") and to have better brainstorming as a result from the situation that every employee can virtually see the data. Therefore, Data Governance can also contribute to innovative approaches in diverse field throughout the whole organization.

3.2.5.2 SMEs

In the case of the SMEs, 57.89% already have experience with Data Governance. Within the observed enterprises, 89.47% of the SMEs do not have an own Data Governance department. On a 10-point Likert scale, the average value for their business value of data is 7.4, whereas the one for data security and data protection is only at 6.78. However, 95% of them have stated that Data Governance can support the appropriate implementation of the GDPR. These results are also different to the interview analysis, which is due to their diverse industries. The less a SME is specialized in a digital or data-driven business, the less Data Governance plays an important role.

The most challenging activities in the context of Data Governance are the

- amount of everyday business that can inhibit an effective Data Governance policy;
- result of a strong bureaucratization;
- difficulty to focus on data security and data protection;
- definition of all related rules, processes and policies;
- difficulty to see data as a value-creating element;
- prospective legal obligations;
- cost intensity; and
- time needed for the implementation.

Some of these results are similar to the general perspective, but also shows that the daily business operations are a particular challenge when introducing Data Governance. This means that in enterprises with a relatively low number of employees, they tend to have a lot of different tasks to work on and therefore have not the chance to focus on a resource-intensive Data Governance strategy.

Further challenges mentioned by the participants are: "Data vs Information vs Process vs Security Responsibilities: CDO vs CIO vs CPO vs CSO" and their integration into the daily workflows.

The most beneficial activities in the context of Data Governance are:

- ensured data security,

- regulatory compliance,
- meeting legal data-related requirements,
- improved customer services and
- having the right data for the right processes.

This also shows a customer-related and legal compliance perspective, which are two main aspects for a SME to stay on the market. This means that with regards to this survey, better decisions due to improved data quality are not as beneficial as the customer orientation itself.

3.2.5.3 Public organizations

The situation of the public organizations having participated in the survey shows that 50% of them already have experience with Data Governance. With only 4 participants for this organizational type, these outcomes are not significant enough. Nevertheless, these results are provided in order to get a rough inside into a few potential forms of dealing with this topic in the respective area.

The most challenging activities in the context of Data Governance are:

- overall increasing amount of data,
- difficulty of a cost-benefit weighing,
- focusing on data security and data protection and
- understanding of Data Governance concepts.

Therefore, the data state that public organizations have to work out the need for a Data Governance strategy in a first step and to communicate these benefits to their decision-makers. The increase in the amounts of data is another problem in this context which might be a severe challenge especially at the beginning. Interestingly, these results are strongly related to some situations in public organizations identified in the interview analysis.

The most beneficial activities in the context of having a Data Governance strategy implemented are:

- better decisions based on better data quality,
- regulatory compliance,

- improved customer services and
- ensured data security.

These opportunities are a mixture of the ones mentioned by the large companies and the SMEs. In addition, they once again reflect most of the statements received during the interviews, and thus show at least some similarities between both surveys of the chosen research method.

4 Summary, conclusions and future work

The increasing presence of a digital business environment permanently enforces companies of all sizes and complexity to rethink their strategy and how to use their data as a critical business resource both internally and externally in an effective way. Thereby, the concept of Data Governance can play an important role. This thesis has followed the goal to provide an insight into the current implementation state of Data Governance as well as resulting challenges and opportunities by providing a systematic literature review and regarding the situation in nine different organizations in Austria. Based on these findings, an online survey was conducted within Austrian and German institutions.

In the light of the nine interviews, the online survey and their analysis, it could be stated that the concept of Data Governance has already found its way into large companies, and also SMEs are concerned with its main issues, especially within the financial and insurance sector as well as in the information and communication industry, although not many of them have had a Data Governance strategy for longer than five years. In this context, public organizations are relatively less involved in this topic and just at the early beginnings of implementing an effective Data Governance strategy.

In general, there are only few challenges which are directly related to a specific organizational type. Far more challenges and opportunities are commonplace for all of them. Moreover, many of these challenges and opportunities supplement each other, which also applies to the relationships of the core dimensions of the DMBOK. External influences such as legal constraints are relatively underrepresented in the survey but have a strong influence. One particular challenge as mentioned in the interviews are the various organizational cultures, especially in large companies. However, its significance has not been confirmed in the online survey.

The majority of participants also sees the potential of Data Governance to positively influence their business and IT strategy. Therefore, it can also en-

able departments to enhance their collaboration, to maximize their potential and to derive new business opportunities from it.

A top-down approach is no necessity, but it is absolutely important to have the commitment from the management when establishing Data Governance. Interestingly, it can be concluded that the size of an organization or heterogeneity of its products is not seen as a limiting factor although it enhances the complexity of such a Data Governance program. Nevertheless, these organizations are facing special difficulties in implementing a Data Governance strategy within their organization due to their complexity and diversity.

The findings have further shown that each organization concerned with Data Governance develops its own model which fits best according to its individual needs since there is no "one size does fit them all" solution. In an ideal case, existing best practices can be considered although they are from mainly American countries and preferably introduced with the help of consultancy firms.

In general, the existing literature focuses on organizational aspects such as how to classify a Data Governance model or to establish a role model. The most common organizational roles, such as data stewards, data experts and data users have been identified in the current state of implementation as well. They are the main actors within a Data Governance program and led by Data Governance coordinators. In several cases it is planned to introduce a Chief Data Officer beside a CIO governing data on the highest executive level.

As for the modification of the DMBOK, organizations have stated that it should be extended by a business perspective since it rather focuses on the IT-related aspects of Data Governance. A possible approach would be to combine this model with the key concepts and principles outlined above. Other dimensions that should be added to this framework are a data processing and a data analytical perspective (such as data science or machine learning). These should support the effective realization of projects with Big Data, Enterprise Linked Data, Data Lakes or Data Spaces, which are emerging trends in enterprise information management. Furthermore, it was argued to interrelate the single dimensions, on the business level with Data Governance, and on the IT-level by using a semantic layer.

Data Governance also has to be truly lived within an organization, i.e. there cannot be a single person responsible for Data Governance alone. Since one particular problem is the lack of sensitivity or information of the employees in the individual departments or perhaps even a false awareness of it, a common feeling for its importance and relevance has to be created.

As for the large companies in the interviews, the overall focus of Data

Governance is rather put on the business aspects. Its implementation ranges from having pilot projects to being too complex to implement within the whole organization. The main driving factors for them have been identified as size, globalization and internal communication. The most challenging aspects were costs estimation, increasing complexity and the heterogeneity of the business and product range, whereas the online survey names cost intensity, legal requirements, identifying data as a value-adding element and data quality issues. On the other hand, particular opportunities are improved business-IT collaboration, better defined rules and responsibilities, data quality, data security and overall data availability and integration.

Overall, the interviews with the SMEs show that the understanding of data Governance concepts is rather diverse. In fact, SMEs do not explicitly use a "real" Data Governance model and focus rather on the technological aspect of Data Governance, which is dependable to the respective industry. Driving factors include the type of business and customers, amount of data and the size of the enterprise. Main challenges which arose were business relevance, measurable benefits, clearly separated roles and its acceptance, whereas the online survey identified difficulties in putting the focus on such a program and contributing the necessary resources to it. Mentioned opportunities include personalized products for customers or the chance to receive certifications from authorities. Beside the improvement in customer relations and availability of data, the better alignment with regulatory requirements has been mentioned.

In case of the public organizations, the implementation state ranges from "not known" to "advanced". Particular challenges provided by the interview partners included its cost intensity, a common understanding of the concept, commitment from the management and having a cultural change. The online survey extended this list by the difficulty to ensure data security and data protection. Related opportunities are better reputation through data quality, less bureaucracy, more efficiency, awareness for data as an asset ("fitness for purpose"), and the implementation of new concepts of data storage and operations. The beneficial aspects from the online survey are better decisions and improved customer services.

However, the validity of the results both from the interviews and the online survey are restricted to the limitations mentioned above due to the observed number of cases and the resulting spread of the results. Since it is a qualitative approach, generalization and objectivity cannot be fulfilled. Moreover, there is no reason why a specific organization is not facing a particular difficulty or seeing a particular opportunity, but was not mentioning it.

In total, it can be concluded that although some challenges and opportunities can be related to a specific type of organization, it is not the case that these are not applicable to other organizational types as well. Consequently, the findings can be used as an orientation but not as an obligation.

As far as future research is concerned, there are still numerous aspects that need to be considered in more detail. First and foremost, since there is no suitable framework for German-speaking countries available, one could use these results for creating a Data Governance framework for a specific organizational type within a particular industry. Second, it should be encouraged to replicate and continue the research process in the future by observing the situation in other countries or industries as well since research in this area is very scarce. For this, the identified challenges and opportunities could represent a valuable basis. Another pivotal research aspect is to survey how Data Governance practices and related challenges and opportunities can change over time with the emergence of new technologies in the respective fields.

The relevance of Data Governance is given by the creation of a value-addressed perspective of data assets by focusing on an organizational collaboration between the business and the IT. Implemented and lived in an appropriate way by regarding respective challenges and opportunities, Data Governance is an important key to business success in any organization. Being able to unlock the value from ones data faster and more efficient than others at any time will likely be a significant competitive advantage. An elaborated Data Governance strategy can be crucial for driving this area forward.
This is a necessary ambition for the future.

5 References

References

[1] Ali M Al-Khouri. Data ownership: who owns "my data". *Int. J. Manag. Inf. Technol*, 2(1):1–8, 2012.

[2] Majid Al-Ruithe, Elhadj Benkhelifa, and Khawar Hameed. Key dimensions for cloud data governance. In *Future Internet of Things and Cloud (FiCloud), 2016 IEEE 4th International Conference on*, pages 379–386. IEEE, 2016.

[3] Wise Analytics. Understanding the role of data governance to support a self-service environment. pages 1–12, 2014.

[4] Data Management Association et al. Dama dmbok functional framework (version 3.02). *DAMA International*, 2008.

[5] Statistics Austria. Classification database. nace rev. 2 - structure. 2017.

[6] Murtha Baca. *Introduction to metadata*. Getty Publications, 2008.

[7] Carolyn Begg and Tom Caira. Exploring the sme quandary: Data governance in practise in the small to medium-sized enterprise sector. *The Electronic Journal Information Systems Evaluation*, 15(1), 2012.

[8] Christian Bizer, Tom Heath, and Tim Berners-Lee. Linked data-the story so far. *Semantic services, interoperability and web applications: emerging concepts*, pages 205–227, 2009.

[9] The Data Protection Officer Blog. The data protection officer. 2017.

[10] Paul Brous, Marijn Janssen, and Paulien Herder. Coordinating data-driven decision-making in public asset management organizations: A quasi-experiment for assessing the impact of data governance on asset management decision making. In *Conference on e-Business, e-Services and e-Society*, pages 573–583. Springer, 2016.

[11] Paul Brous, Marijn Janssen, and Riikka Vilminko-Heikkinen. Coordinating decision-making in data management activities: A systematic review of data governance principles. In *International Conference on Electronic Government and the Information Systems Perspective*, pages 115–125. Springer, 2016.

[12] Erik Brynjolfsson, Lorin M Hitt, and Heekyung Hellen Kim. Strength in numbers: How does data-driven decisionmaking affect firm performance? 2011.

[13] Information Builders. Seven steps to effective data governance. pages 1–16, 2011.

[14] Capgemini. Data governance for financial institutions. pages 1–12, 2013.

[15] Lai Kuan Cheong and Vanessa Chang. The need for data governance: a case study. *ACIS 2007 Proceedings*, page 100, 2007.

[16] David W Coleman, Allen A Hughes, and Wayne D Perry. The role of data governance to relieve information sharing impairments in the federal government. In *Computer Science and Information Engineering, 2009 WRI World Congress on*, volume 4, pages 267–271. IEEE, 2009.

[17] European Commission. Official journal of the european communities (l 195). *Official Journal of the European Union (OJ)*, 23:35–37, 1980.

[18] IBM Corporation. Successful information governance through high-quality data. somers. pages 1–12, 2012.

[19] IBM Data Governance Council Maturity Council, IDG. The ibm data governance council maturity model: Building a roadmap for effective data governance, 2007.

[20] P Cupola, S Earley, and D Henderson. Dama-dmbok2 framework, 2014.

[21] DataFlux. The data governance maturity model. establishing the people, policies and technology that manage enterprise data. pages 1–11, 2007.

[22] Fernandez Garcia Javier David, Elmar Kiesling, Sabrina Kirrane, Julia Neuschmid, Nika Mizerski, Axel Polleres, Marta Sabou, Thomas Thurner, and Peter Wetz. Propelling the potential of enterprise linked data in austria. roadmap and report. pages 1–200, 2016.

[23] Magistratsdirektion der Stadt Wien. Data excellence in der stadt wien. pages 1–25, 2017.

[24] RJ DeStefano, Lixin Tao, and Keke Gai. Improving data governance in large organizations through ontology and linked data. In *Cyber Security and Cloud Computing (CSCloud), 2016 IEEE 3rd International Conference on*, pages 279–284. IEEE, 2016.

[25] J Dyché. A data governance manifesto: Designing and deploying sustainable data governance. *Baseline Consulting, Los Angeles, California*, 2007.

[26] Andrea Fontana and James H Frey. The interview: From structured questions to negotiated text. *Handbook of qualitative research*, 2(6):645–672, 2000.

[27] Michael Franklin, Alon Halevy, and David Maier. From databases to dataspaces: a new abstraction for information management. *ACM Sigmod Record*, 34(4):27–33, 2005.

[28] Ed Gelbstein. Is audit basics: Data management body of knowledge - a summary for auditors. 2017.

[29] Niamh Gleeson and Ian Walden. Placing the state in the cloud: Issues of data governance and public procurement. *Computer Law & Security Review*, 32(5):683–695, 2016.

[30] Adrian Gregory. Data governance-protecting and unleashing the value of your customer data assets. *Journal of Direct, Data and Digital Marketing Practice*, 12(3):230–248, 2011.

[31] Jane Griffin. Four critical principles of data governance success. *Information Management*, 20(1):28, 2010.

[32] Entity Group. Is effective data governance a choice or a necessity in financial services today. pages 1–8, 2014.

[33] Malcolm W Harkins. Managing risk and information security.

[34] Anna Horlacher and Thomas Hess. What does a chief digital officer do? managerial tasks and roles of a new c-level position in the context of digital transformation. In *System Sciences (HICSS), 2016 49th Hawaii International Conference on*, pages 5126–5135. IEEE, 2016.

[35] George Hripcsak, Meryl Bloomrosen, Patti FlatelyBrennan, Christopher G Chute, Jim Cimino, Don E Detmer, Margo Edmunds, Peter J Embi, Melissa M Goldstein, William Ed Hammond, et al. Health data use, stewardship, and governance: ongoing gaps and challenges: a report from amia's 2012 health policy meeting. *Journal of the American Medical Informatics Association*, 21(2):204–211, 2014.

[36] Jessica Claudia Iacono, Ann Brown, and Clive Holtham. The use of the case study method in theory testing: The example of steel emarketplaces. *Electronic Journal of Business Research Methods*, 9(1), 2011.

[37] The Data Governance Institute. Definitions of data governance. 2015.

[38] Rob Karel. Data governance: What works and what doesn't. *Forrester Research*, 2007.

[39] Wolfgang Kerber. Governance of data: Exclusive property vs. access, 2016.

[40] Vijay Khatri and Carol V Brown. Designing data governance. *Communications of the ACM*, 53(1):148–152, 2010.

[41] Tibor Koltay. Data governance, data literacy and the management of data quality. *IFLA Journal*, 42(4):303–312, 2016.

[42] John Ladley. *Data Governance: How to Design, Deploy, and Sustain an Effective Data Governance Program*. Newnes, 2012.

[43] N Laskowski. Data lake governance: A big data do or die. 2016.

[44] Jerry N Luftman, Christine V Bullen, Donald Liao, Elby Nash, and Carl Neumann. *Managing the information technology resource: Leadership in the information age*. Pearson Education Upper Saddle River, 2004.

[45] Cedrine Madera and Anne Laurent. The next information architecture evolution: the data lake wave. In *Proceedings of the 8th International Conference on Management of Digital EcoSystems*, pages 174–180. ACM, 2016.

[46] Natalia Miloslavskaya and Alexander Tolstoy. Big data, fast data and data lake concepts. *Procedia Computer Science*, 88:300–305, 2016.

[47] Vincenzo Morabito. Big data governance. In *Big data and analytics*, pages 83–104. Springer, 2015.

[48] Michael D Myers and Michael Newman. The qualitative interview in is research: Examining the craft. *Information and organization*, 17(1):2–26, 2007.

[49] Valeh Nazemoff. Whitepaper data governance roadmap for it executives. pages 1–14, 2010.

[50] Javier Nogueras-Iso, F Javier Zarazaga-Soria, Javier Lacasta, Rubén Béjar, and Pedro R Muro-Medrano. Metadata standard interoperability: application in the geographic information domain. *Computers, environment and urban systems*, 28(6):611–634, 2004.

[51] Boris Otto. Data governance. *Business & Information Systems Engineering*, 3(4):241–244, 2011.

[52] Boris Otto. A morphology of the organisation of data governance. In *ECIS*, volume 20, page 1, 2011.

[53] Boris Otto. Organizing data governance: findings from the telecommunications industry and consequences for large service providers. 2011.

[54] Boris Otto. On the evolution of data governance in firms: the case of johnson & johnson consumer products north america. In *Handbook of Data Quality*, pages 93–118. Springer, 2013.

[55] Boris Otto and Kristin Weber. Data governance. In *Daten-und Informationsqualität*, pages 269–286. Springer, 2015.

[56] Zeljko Panian. Some practical experiences in data governance. *World Acad. Sci. Eng. Technol*, 38:150–157, 2010.

[57] E Pierce, WS Dismute, and C Lwanga Yonke. The state of information and data governance–understanding how organizations govern their information and data assets. *IAIDQ and UALR-IQ*, 2008.

[58] Dan Power. Data governance: Start from where you are. pages 1–11, 2012.

[59] Torsten Priebe and Stefan Markus. Business information modeling: A methodology for data-intensive projects, data science and big data governance. In *Big Data (Big Data), 2015 IEEE International Conference on*, pages 2056–2065. IEEE, 2015.

[60] Protiviti. Data governance combined with analytics for powerful business insights. pages 1–3, 2016.

[61] PwC. Data governance survey results: A european comparison of data management capabilities in banks. pages 1–36, 2016.

[62] Mohammad Rifaie, Reda Alhajj, and Mick Ridley. Data governance strategy: A key issue in building enterprise data warehouse. In *Proceedings of the 11th International Conference on Information Integration and Web-based Applications & Services*, pages 587–591. ACM, 2009.

[63] Frederick J Riggins and Bonnie K Klamm. Data governance case at krausemcmahon llp in an era of self-service bi and big data. *Journal of Accounting Education*, 38:23–36, 2017.

[64] Sara Rosenbaum. Data governance and stewardship: designing data stewardship entities and advancing data access. *Health services research*, 45(5p2):1442–1455, 2010.

[65] Javier Salido and P Voon. A guide to data governance for privacy, confidentiality, and compliance. *Microsoft Trust. Comput.*, page 35, 2010.

[66] Sunil Soares. *Big data governance: An emerging imperative*. Mc Press, 2012.

[67] Helen Sun. Enterprise information management: Best practices in data governance. pages 1–20, 2011.

[68] Paul P Tallon. Corporate governance of big data: Perspectives on value, risk, and cost. *Computer*, 46(6):32–38, 2013.

[69] MapR Technologies. The need for big data governance a whitepaper by collibra and mapr. pages 1–8, 2015.

[70] Gwen Thomas. *Alpha males and data disasters: The case for data governance*. Brass Cannon Press Orlando, FL, 2006.

[71] Gwen Thomas. The dgi data governance framework. *The Data Governance Institute, Orlando, FL (USA)*, 2012.

[72] Nik Thompson, Ravi Ravindran, and Salvatore Nicosia. Government data does not mean data governance: Lessons learned from a public sector application audit. *Government information quarterly*, 32(3):316–322, 2015.

[73] Roland L Trope, E Michael Power, Vincent I Polley, and Bradford C Morley. A coherent strategy for data security through data governance. *IEEE Security & Privacy*, 5(3), 2007.

[74] Wil MP Van der Aalst. Data scientist: The engineer of the future. In *Enterprise Interoperability VI*, pages 13–26. Springer, 2014.

[75] Stefhan van Helvoirt and Hans Weigand. Operationalizing data governance via multi-level metadata management. In *Conference on e-Business, e-Services and e-Society*, pages 160–172. Springer, 2015.

[76] Günter Verheugen. The new sme definition: user guide and model declaration. *Enterprise and Industry Publications, European Commission*, 2005.

[77] Richard Y Wang and Diane M Strong. Beyond accuracy: What data quality means to data consumers. *Journal of management information systems*, 12(4):5–33, 1996.

[78] Kristin Weber, Boris Otto, and Hubert Österle. Data governance: Organisationskonzept für das konzernweite datenqualitätsmanagement. 2009.

[79] Kristin Weber, Boris Otto, and Hubert Österle. One size does not fit all—a contingency approach to data governance. *Journal of Data and Information Quality (JDIQ)*, 1(1):4, 2009.

[80] Nicholas M Weber, Carole L Palmer, and Tiffany C Chao. Current trends and future directions in data curation research and education. *Journal of Web Librarianship*, 6(4):305–320, 2012.

[81] Kristin Wende. A model for data governance-organising accountabilities for data quality management. 2007.

[82] Kristin Wende and Boris Otto. A contingency approach to data governance. 2007.

[83] David Wood, Marsha Zaidman, Luke Ruth, and Michael Hausenblas. *Linked Data*. Manning Publications Co., 2014.

6 Appendix

6.1 Additional tables

The following section contains all the tables that have been either too long for the main part of the thesis or have been identified as only representing additional information as part of the creation of the thesis or background knowledge.

Table 5: Summary of the literature review (scientific literature and white papers).

Date	Database	Search terms	Relevant period	Screened	Appropriate
June 15, 2017	Google Scholar	"data governance"	All	70	15
June 15, 2017	Google Scholar	"data governance challenges"	All	15	0
June 15, 2017	Google Scholar	"data governance opportunities"	All	15	0
June 15, 2017	Google Scholar	"data governance strategy"	All	11	2
June 15, 2017	Google Scholar	"data governance model/roles"	All	20	9
June 15, 2017	Google Scholar	"data governance best practices"	All	9	1
June 15, 2017	Google Scholar	"data governance public"	All	16	4
June 15, 2017	Google Scholar	"data governance company"	All	18	3
June 15, 2017	Google Scholar	"data governance sme"	All	12	1
June 15, 2017	Google Scholar	"data lake"	All	9	3
June 15, 2017	Google Scholar	"data space"	All	6	1
June 15, 2017	Google Scholar	"enterprise linked data"	All	7	4
June 15, 2017	Google Scholar	"big data"	All	16	4
June 15, 2017	Google Scholar	"data value organization"	All	15	7
June 15, 2017	Google Scholar	"chief data officer"	All	5	1
June 15, 2017	Google Scholar	"data governance security/privacy"	All	10	3
June 15, 2017	Google Scholar	"data scientist"	All	6	1
June 15, 2017	Google Scholar	"data governance white paper"	All	41	13
Sum				311	72

Table 6: Definition of Data Governance, separated by identified elements.

	Reference											
Element	1	2	3	4	5	6	7	8	9	10	11	12
Specify framework for decision rights and accountabilities	[79]	[70]				[4]	[10]	[30]		[40]	[55]	[82]
Encourage desirable behavior in the use of data	[79]			[11]						[40]	[55]	[82]
Develop and implement corporate-wide data policies, guidelines, and standards that are consistent with the organization's mission, strategy, values, norms, and culture	[79]		[60]					[30]				

Manage the quantity, consistency, usability, security and availability of data (i.e. data quality)		[70]		[11]	[38]	[4]	[10]	[30]	[56]			
Collection of processes		[70]	[60]					[30]				
Standards, policies			[60]						[56]			
Technologies								[30]	[56]			
Strategic business program					[38]			[30]				
Planning, supervision and control over data management and use									[56]			

Table 7: The updated version of the DMBOK model, based on the DAMA-DMBOK Functional Framework [20].

		Process	Technology	People	Technology	Process	People
Dimension	Goals & Principles	Activities	Deliverables	Roles & Responsibilities	Tools	Practices & Techniques	Organization & Culture
Data Quality							
Metadata							
Data Warehousing & Business Intelligence							
Reference & Master Data							
Documents & Content Management							
Data Architecture							
Data Modeling & Design							
Data Storage & Operations							
Data Security							
Data Integration & Interoperability							
Data Governance							

Table 8: Summary of challenges of Data Governance mentioned in the interviews and surveyed with the online questionnaire.

No.	Challenge	Reference	Interview	Organization	Type of challenge	Dimension	Mean	Large	SMEs	Public
1	Understanding of DG concepts (e.g. definition, certain dimension or idea)	[47], [33]	6/9	Large company, SME, public organization	Knowledge and experience	Internal	5.16	5.81	4.47	5.50
2	Complexity of a DG Framework	[33], [56]	3/9	Large company, public organization	Knowledge and experience	Internal	6.24	6.88	4.88	4.75
3	Requirement of information, education, training	[82], [10]	3/9	Large company, public organization	Knowledge and experience	Internal	6.00	6.24	5.35	5.00
4	Lack of best practices	[67], [25]	3/9	Large company, public organization	Knowledge and experience	External	6.48	6.71	5.18	4.75
5	Implementation takes a long time	[47]	5/9	Large company, SME, public organization	Organizational and cultural	Internal	6.62	6.94	5.65	4.00
6	Size of the organization	[68]	5/9	Large company, SME, public organization	Organizational and cultural	Internal	5.24	6.06	3.65	4.75

No.	Challenge	Reference	Interview	Organization	Type of challenge	Dimension	Mean	Large	SMEs	Public
7	Having a cultural change within the organization	[30]	4/9	Large company, SME, public organization	Organizational and cultural	Internal	6.14	6.53	5.06	4.25
8	Finding a suitable maturity model	[58]	4/9	Large company, SME, public organization	Organizational and cultural	Internal	5.40	6.25	4.71	4.00
9	Definition of all related rules, processes and policies	[40]	3/9	Large company, public organization	Organizational and cultural	Internal	6.81	7.35	5.71	5.25
10	Has not been pursued well enough until now	[35]	3/9	Large company, public organization	Organizational and cultural	Internal	5.75	5.94	4.94	4.50
11	Fear coming along with such a change	[47]	3/9	Large company, public organization	Organizational and cultural	Internal	5.33	6.06	3.06	4.25
12	Collaboration with a consulting firm	[63]	3/9	Large company, public organization	Organizational and cultural	Internal	4.19	4.65	2.76	3.75
13	Heterogeneity of the business and diverse product ranges	[68]	3/9	Large company, SME	Organizational and cultural	Internal	4.95	4.94	4.41	4.25

No.	Challenge	Reference	Interview	Organization	Type of challenge	Dimension	Mean	Large	SMEs	Public
14	Too much everyday business inhibits an effective DG policy	[40]	2/9	Large company, public organization	Organizational and cultural	Internal	6.00	5.47	6.18	4.25
15	Implicit and fragmentary implementation inhibits an explicit overall implementation	[33]	2/9	SME, public organization	Organizational and cultural	Internal	5.05	5.29	4.75	4.75
16	Can lead to a strong bureaucratization	[60]	1/9	Large company	Organizational and cultural	Internal	6.19	6.06	5.94	5.50
17	Overall increasing amount of data	[68]	9/9	Large company, SME, public organization	Strategic and business	Internal	6.33	6.29	6.12	5.50
18	Seeing data as a value-creating element	[51]	9/9	Large company, SME, public organization	Strategic and business	Internal	6.10	6.94	5.69	4.75
19	Top-down or bottom-up approach	[47]	8/9	Large company, SME, public organization	Strategic and business	Internal	5.19	6.18	4.41	3.50
20	Necessity for new employees	[70], [82]	2/9	Public organization	Strategic and business	Internal	5.10	5.47	3.88	4.75

No.	Challenge	Reference	Interview	Organization	Type of challenge	Dimension	Mean	Large	SMEs	Public
21	Difficulty of a cost-benefit weighing	[30]	8/9	Large company, SME, public organization	Monetary	Internal	5.62	6.00	5.18	5.50
22	Estimation of the costs involved	[30]	8/9	Large company, SME, public organization	Monetary	Internal	5.52	5.88	4.82	4.25
23	Cost intensity	[30]	6/9	Large company, SME, public organization	Monetary	Internal	7.05	7.71	5.65	5.00
24	Effort necessary to ensure data quality	[52]	7/9	Large company, SME, public organization	Data quality	Internal	6.43	6.82	5.24	5.00
25	Focusing on data security and data protection	[47], [69]	6/9	Large company, SME, public organization	Security and privacy	Internal	6.19	6.35	5.76	5.50
26	Prospective legal obligations	[38]	5/9	Large company, SME, public organization	Legal	External	6.25	7.12	5.69	4.50
27	Dealing with data-related challenges from the past	[68]	4/9	Large company, public organization	Technical	Internal	5.43	5.71	5.50	3.75

Table 9: Summary of challenges of Data Governance mentioned in the interviews.

No.	Challenge	Reference	Interview	Organization	Type of challenge	Dimension
28	Underestimating the benefits of Data Governance	[30], [68]	8/9	Public organization	Knowledge and experience	Internal
29	Knowhow to evaluate these amounts of data	[60]	7/9	Large company, SME, public organization	Knowledge and experience	Internal
30	Making data quality measurable since it is very difficult to define	[41]	7/9	Large company, SME, public organization	Knowledge and experience	Internal
31	Unawareness of the topic since it is not widely known yet	[15]	6/9	Large company, SME, public organization	Knowledge and experience	Internal
32	Understand the needs and benefits of the realization	[33], [82]	5/9	Large company, public organization	Knowledge and experience	Internal
33	Experience within the company required for introducing a change	[10], [82]	5/9	Large company, public organization	Knowledge and experience	Internal

No.	Challenge	Reference	Interview	Organization	Type of challenge	Dimension
34	Need for detailed work with the concept to understand its potential opportunities	[10], [82]	5/9	Large company, public organization	Knowledge and experience	Internal
35	Different frameworks for different industries	[79],[15]	1/9	SME	Knowledge and experience	Internal
36	Complexity of the organization	[32]	7/9	Large company, SME, public organization	Organizational and cultural	Internal
37	Convince the employees about necessity	[51]	6/9	Large company, SME, public organization	Organizational and cultural	Internal
38	Resistance against the introduction of new organizational concepts	[47]	6/9	Large company, SME, public organization	Organizational and cultural	Internal

No.	Challenge	Reference	Interview	Organization	Type of challenge	Dimension
39	Many different requirements of different departments or sub-divisions within an organization	[52], [59]	6/9	Large company, public organization	Organizational and cultural	Internal
40	Communication & implementation is difficult to manage across departments	[79]	5/9	Large company, public organization	Organizational and cultural	Internal
41	Does not stop, it is a continuous process	[30]	4/9	Large company, SME, public organization	Organizational and cultural	Internal
42	Unclear division of roles	[40]	4/9	Large company, public organization	Organizational and cultural	Internal
43	Slow progress due to lack of resources and professional personnel	[40]	3/9	Large company, public organization	Organizational and cultural	Internal
44	Speed of appearance of new amounts of data	[72]	9/9	Large company, SME, public organization	Strategic and business	External

No.	Challenge	Reference	Interview	Organization	Type of challenge	Dimension
45	Bad data quality results in worse reputation within the people or customers	[41]	8/9	Large company, SME, public organization	Strategic and business	External
46	Inability to implement with the help of technology alone	[21]	8/9	Large company, SME, public organization	Strategic and business	Internal
47	Support by the management required	[56]	6/9	Large company, SME, Public organization	Strategic and business	Internal
48	Finding the appropriate use & implementation is hard	[21]	6/9	Large company, SME, public organization	Strategic and business	Internal
49	Drawing conclusions from the data through Data Science or Machine Learning	[60], [69]	6/9	Large company, SME, public organization	Strategic and business	Internal
50	Serving the entire process with data	[41]	5/9	Large company, SME, public organization	Strategic and business	Internal

No.	Challenge	Reference	Interview	Organization	Type of challenge	Dimension
51	Creation of a detailed business object model	[21]	4/9	SME, public organization	Technical	Internal

Table 10: Summary of opportunities of Data Governance mentioned in the interviews and surveyed with the online questionnaire.

No.	Opportunity	Reference	Interview	Organization	Type of challenge	Dimension	Mean	Large	SMEs	Public
1	Having Data Scientists	[74]	8/9	Large company, SME, public organization	Organizational and cultural	Internal	5.14	5.81	4.07	4.25
2	Work-supportive function instead of being a barrier	[10]	6/9	Large company, SME, public organization	Organizational and cultural	Internal	6.52	6.82	5.94	4.25
3	Organizational roles are distributed more clearly	[56], [70], [15]	4/9	Large company, SME, public organization	Organizational and cultural	Internal	6.05	6.18	5.25	5.25
4	Treating data as a valuable business asset	[51]	9/9	Large company, SME, public organization	Strategic and business	Internal	7.33	7.59	7.00	4.50
5	Bridging the gap between the IT and the business	[68]	9/9	Large company, SME, public organization	Strategic and business	Internal	6.57	6.41	6.06	5.00
6	Data Lakes or Big Data projects can be better implemented	[66], [46]	8/9	Large company, SME, public organization	Strategic and business	Internal	5.48	6.06	4.56	4.75

No.	Opportunity	Reference	Interview	Organization	Type of challenge	Dimension	Mean	Large	SMEs	Public
7	Having the right data for the right processes	[41]	8/9	Large company, SME, public organization	Strategic and business	Internal	7.30	7.71	6.60	4.50
8	Better decisions based on better data quality	[52], [41]	7/9	Large company, SME, public organization	Strategic and business	Internal	7.43	8.24	6.19	6.00
9	Efficiencies in business processes	[31], [35], [42]	6/9	Large company, SME, public organization	Strategic and business	Internal	6.43	6.82	6.00	4.75
10	Improved customer services	[14]	5/9	Large company, SME, public organization	Strategic and business	External	6.71	6.88	6.75	5.75
11	Strengthen the external confidence with the organization	[7]	5/9	Large company, SME, public organization	Strategic and business	External	6.10	6.24	5.69	4.50
12	Reduced customer complaints	[51], [52]	5/9	Large company, SME, public organization	Strategic and business	External	5.10	5.88	5.07	5.00
13	Data-related processes are covered by responsibilities	[70], [15]	5/9	Large company, SME, public organization	Strategic and business	Internal	6.71	7.19	6.13	5.25

No.	Opportunity	Reference	Interview	Organization	Type of challenge	Dimension	Mean	Large	SMEs	Public
14	Having pilot projects	[10]	4/9	Large company, SME, public organization	Strategic and business	Internal	5.14	5.94	4.13	4.25
15	Enabling more efficient gain of business related information and knowledge	[50]	4/9	Large company, public organization	Strategic and business	Internal	6.75	7.06	6.13	5.00
16	Having various Key Performance Indicators to measure the success of DG	[60], [14]	1/9	Public organization	Strategic and business	Internal	5.33	6.18	4.25	5.00
17	Less waste of resources	[53]	1/9	SME	Strategic and business	Internal	5.71	6.24	5.44	3.75
18	Enabling personalized products for customers	[14]	1/9	SME	Strategic and business	External	6.10	7.06	4.94	4.75
19	Reduced transaction costs	[22]	2/9	Large company, SME	Monetary	Internal	4.85	5.71	4.53	3.75
20	Gaining value from prospective data markets	[65]	2/9	SME, Public organization	Monetary	External	6.43	6.82	5.87	5.50

No.	Opportunity	Reference	Interview	Organization	Type of challenge	Dimension	Mean	Large	SMEs	Public
21	Ensured data security	[65], [40], [56], [72]	5/9	Large company, SME, public organization	Security and privacy	Internal	7.95	8.06	7.50	5.75
22	Consistent data quality policies	[52], [41]	8/9	Large company, SME, public organization	Data quality	Internal	6.90	7.29	6.00	5.25
23	Meeting legal data-related requirements	[68], [66], [56]	7/9	Large company, SME, public organization	Legal	External	7.71	7.76	7.38	5.25
24	Regulatory compliance	[68], [66], [56]	7/9	Large company, SME, public organization	Legal	External	7.00	7.29	7.44	5.75
25	Simplified access to relevant data	[15]	6/9	Large company, SME, public organization	Technical	Internal	6.62	7.06	6.19	4.00
26	Unifying various data-related processes	[7]	6/9	Large company, SME, public organization	Technical	Internal	6.48	6.88	6.00	4.25
27	Minimized data redundancies	[21]	5/9	Large company, SME, public organization	Technical	Internal	6.81	7.47	6.00	3.25

Table 11: Summary of opportunities of Data Governance mentioned in the interviews.

No.	Opportunity	Reference	Interview	Organization	Type of challenge	Dimension
28	Less separation in between the different departments and their data	[3], [49]	7/9	Large company, SME, public organization	Organizational and cultural	Internal
29	All data-related processes are completely covered by responsibilities	[70], [15]	5/9	Large company, SME, public organization	Organizational and cultural	Internal
30	Having a unique specific framework for the organization	[79]	4/9	Large company, SME, public organization	Organizational and cultural	Internal
31	Internal media and small reports	N/A	1/9	Public organization	Organizational and cultural	Internal
32	Implement a Data Governance House which tries to summarize best practices	N/A	1/9	Public organization	Organizational and cultural	Internal

No.	Opportunity	Reference	Interview	Organization	Type of challenge	Dimension
33	Quality of staff and data are both important	[79]	9/9	Large company, SME, public organization	Strategic and business	Internal
34	Constant availability of up-to-date data where necessary	[11]	5/9	Large company, SME, public organization	Strategic and business	Internal
35	Making data-related processes more consistent, transparent and manageable	[56], [11]	4/9	Large company, SME, public organization	Strategic and business	Internal
36	Having described a suitable data model	[21]	3/9	Large company, public organization	Strategic and business	Internal
37	Appropriate architecture and planning allow high speed operations	[52], [21]	3/9	SME, public organization	Strategic and business	Internal

No.	Opportunity	Reference	Interview	Organization	Type of challenge	Dimension
38	Available data are used fully and appropriately in the entire data value creation chain	[68]	2/9	SME, public organization	Strategic and business	Internal
39	Being implementable in any industry	[68]	1/9	SME	Strategic and business	Internal
40	Managing a mix of complex products	[68]	1/9	SME	Strategic and business	Internal
41	Ensuring certifications from authorities	N/A	1/9	SME	Strategic and business	External
42	Clear data origin due to data integrity	[6], [75]	5/9	Large company, public organization	Technical	Internal
43	Business Intelligence works better if the data are integrated	[63], [38]	4/9	SME, public organization	Technical	Internal
44	Having correct data sets	N/A	2/9	SME	Technical	Internal

No.	Opportunity	Reference	Interview	Organization	Type of challenge	Dimension
45	Having an Enterprise Architecture Management (EAM) tool	N/A	1/9	Large company	Technical	Internal

6.2 Transcripts of the interviews

The following section should have shown the transcripts of the interviews for which were given the permission by the interview partners to publish them at the end of this thesis. It should be mentioned that almost all of the interview partners supposed to have the transcript in this appendix after the interview. However, all nine interview partners had to dismiss after having a discussion about it within their specific organization since the topic is very sensitive and should therefore not be published. In fact, it would have even been relatively easy to conclude from an anonymous version which concrete organization is concerned. As a results, this section remains empty.

6.3 Blank interview questionnaire

The interview guideline that follows represents the total amount of questions taken into account for all nine interviews which were conducted. Its diversity has been chosen due to the uncertain specific situation in all the different companies in order to understand these differences as precise as possible. Each Interview started with a short presentation of the interviewer, followed by the semi-structured interviews.

0. **Vorstellung der Interviewpartners**
    a. Was sind Ihre Tätigkeiten und Zuständigkeiten im Unternehmen?
    b. Wie lange arbeiten Sie schon in diesem Bereich?

1. **Definition, Begriffsverständnis und Framework-Analyse**
    a. Zunächst einmal stellt sich die Frage, inwieweit Sie und Ihr Unternehmen mit dem Begriff der Data Governance vertraut sind. Wie würden Sie den Begriff der Data Governance – basierend auf Ihren betrieblichen Erfahrungen – (für Ihr Unternehmen) definieren?
    b. Was sind Ihre Erfahrungen mit Data Governance?
    c. Als nächstes möchte ich Ihnen bitte einen Framework vorstellen, der für die Analyse verwendet wird: dieser Data Management Body of Knowledge (DMBOK) besteht aus insgesamt 10 verschiedenen Dimensionen, die sich wiederum in mehrere Unterpunkte aufteilen. Wenn Sie diese Grafik betrachten, welche dieser Dimensionen sind Ihnen bereits bekannt und werden in Ihrem Unternehmen umgesetzt?
    d. Welche davon gehören in Ihr Verständnis von Data Governance?
    e. Welches Modell bzw. Framework verwenden Sie in Ihrem Unternehmen?
    f. Haben Sie sich selbst eines entwickelt oder sind Sie gerade dabei, eines einzuführen?
    g. Greifen Sie auf Best Practices zurück? Wenn ja, auf welche?

2. **Vertiefende Betrachtung des DMBOK Frameworks**
    a. Bitte gehen Sie auf zwei dieser Dimensionen etwas näher ein. Wie haben Sie diese bereits implementiert?
    b. Wie würden Sie deren Wichtigkeit in absteigender Reihenfolge definieren?

3. **Bedeutung und Einführung der Data Governance**
    a. Wofür benötigen Sie im Unternehmen Data Governance?
    b. Wie ist es um Ihre aktuelle Data Governance bestellt?
    c. Wie, wann und warum hat die Implementierung von Data Governance in Ihrem Unternehmen begonnen?
    d. Was macht eine gute Data Governance aus? Was zeichnet Ihre Data Governance aus?
    e. Gibt es Bereiche von Data Governance, die bereits fortgeschrittener als andere sind?
    f. Wie lange dauert es, Data Governance einzuführen?
    g. Beginnen Sie dabei in einer bestimmten Abteilung bzw. einem bestimmten Bereich?

4. **Zuständigkeiten und Rollen**
    a. Welche Rollen sind daran beteiligt?
    b. Wie werden diese Rollen bzw. Verantwortlichkeiten untereinander festgelegt?
    c. Wer verantwortet Data Governance in Ihrem Unternehmen?
    d. Brauchen Sie dafür zusätzliches Personal?
    e. Welche Rolle spielen Data Scientists?
    f. Haben Sie eine eigene Data Governance Abteilung?
    g. Gibt es in Ihrem Unternehmen bereits einen eigene Chief Data Officer (CDO)? Wenn nein, wer erfüllt dann diese Aufgaben?

5. **Kommunikation und Umsetzung**
    a. Viele Unternehmen scheinen sich mit abteilungsbezogener Data Governance zu beschäftigen. Wie könnte aus Ihrer Sicht der Umstieg auf eine Unternehmensinitiative (schneller/besser) erreicht werden?
    b. Wie wird mit organisationsinternen Konflikten sich überlappender Funktionen bei der Handhabung von Daten umgegangen?
    c. Wie ist die Konversation zwischen Management und den zuständigen Mitarbeitern für die Umsetzung der Data Governance in Ihrem Unternehmen organisiert?
    d. Arbeiten Sie über die einzelnen Abteilungen hinweg um Data Governance umzusetzen?

6. **Bedeutung und Wert von Daten**
    a. Wie würden Sie die Bedeutung von Daten für Ihr Unternehmen einordnen (ist z.B. das Personal wichtiger als die Daten)?
    b. Wovon ist der Wert Ihrer Daten abhängig?
    c. Was kostet eine mögliche schlechte Datenqualität für Ihr Unternehmen?
    d. Wie werden Daten in Ihrem Unternehmen verarbeitet?

7. **Datensicherheit und Datenschutz**
    a. Gerade in Bezug auf Datensicherheit – einer wichtigen Dimension der Data Governance – kommt ab Mai 2018 eine große Veränderung auf alle Unternehmen zu. Bis dahin muss nämlich die Datenschutzgrundverordnung (DS-GVO) der EU umgesetzt sein. Sind Sie bereits ausreichend darauf vorbereitet?
    b. Wie war es im Vergleich dazu zu Zeiten des Datenschutzgesetzes (2000)?
    c. Wie gehen Sie mit sensiblen Daten um?
    d. Wie sind Datensicherheit- und Datenschutz in Ihrem Unternehmen umgesetzt? Welche Rollen existieren dafür?

8. **Allgemeine Chancen und Herausforderungen von Data Governance**
    a. Was sind Ihrer Meinung nach die großen Herausforderungen und Chancen für Ihr Unternehmen, die mit Data Governance einhergehen?
    b. Wie würden Sie einzelne Chancen und Herausforderungen, die sich durch Data Governance ergeben, beurteilen?
    c. Wie würden Sie den Erfolg von Data Governance messen?
    d. Welche zusätzlichen Chancen und Herausforderungen sehen Sie für Ihr Unternehmen?
    e. Welche Möglichkeiten erhoffen Sie sich dadurch? Welche Konsequenzen hat das?
    f. Welche Probleme haben Sie bei der Einführung, Umsetzung, Weiterentwicklung etc. dieser Maßnahmen erlebt?

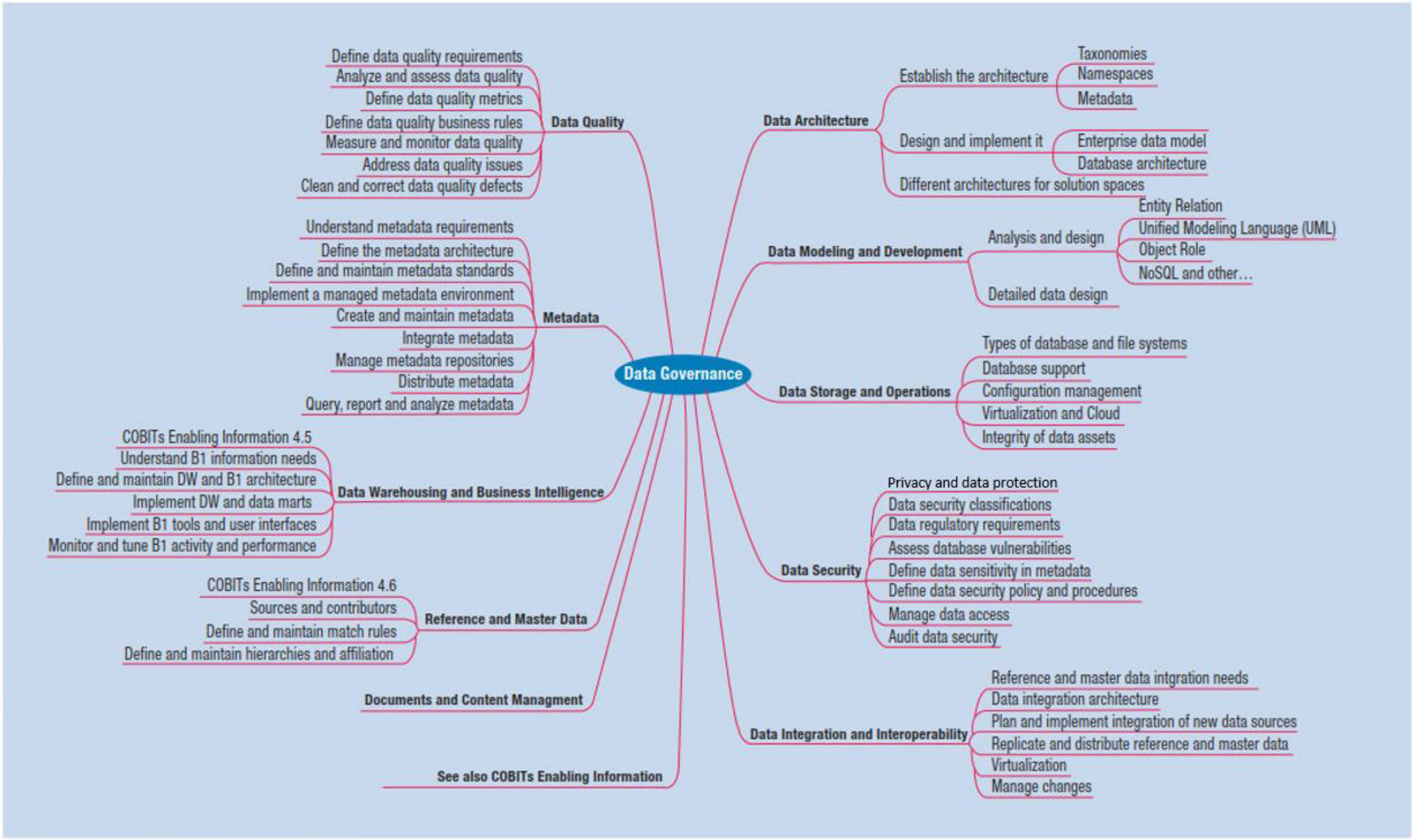
9. **Data Lakes, Data Spaces, Enterprise Linked Data, Big Data**
    a. Sind Konzepte wie Data Lakes, Data Spaces, Enterprise Linked Data und Big Data bereits in Ihrem Unternehmen angekommen?
    b. Was sind die größten Herausforderungen und Chancen im Kontext von Big Data, Data Lakes, Enterprise Linked Data und Data Spaces?
    c. Wie managt Ihr Unternehmen seine Daten und durch welche neuen Einflüsse (Big Data, Data Lakes etc.) sind diese bestimmt worden?
    d. Finden Sie, dass es sinnvoll ist Data Governance auf Data Lakes anzuwenden oder nur auf Daten die bereits für einen bestimmten Geschäftszweck herausgeholt wurden?
    e. Wie werden die Daten in Ihrem Unternehmen gelagert (heute und in der Vergangenheit)?
    f. Was hat bei Ihnen in der Vergangenheit gut funktioniert und warum ist dies vielleicht heute nicht mehr möglich?
    g. Wie ist der Zugriff auf die Daten geregelt?
    h. Welche neuen Technologien und Rollen hat Ihr Unternehmen dafür bereitgestellt?
    i. Verwenden Sie Systeme, die die Bereitstellung von Datenqualität garantieren und das Risiko reduzieren?
    j. Welche neuen Technologien und bezogenen Infrastruktur waren/sind notwendig zu implementieren um ein unternehmensweites Data Governance Programm zu unterstützen?

10. Gesamtbetrachtung, Reifegrad, strategische Bedeutung und Ziele

a. Ist Data Governance für Ihr Unternehmen eher ein IT, ein business program, oder beides?
b. Welche Rolle haben Daten in Ihrer Corporate Strategy?
c. Haben diese überhaupt eine strategische Rolle oder sind sie nur „verpflichtend“ zu verwalten?
d. Würden Sie sagen dass die Bedeutung von Data Governance derzeit noch unterschätzt wird?
e. Wo würden Sie sich im Vergleich mit Ihren Mitbewerbern in Hinblick auf Ihre Data Governance positionieren?
f. Wie würden Sie den aktuellen Reifegrad Ihrer Data Governance (Data Governance Maturity) auf einer Skala von 1-5 einstufen?
g. Wie beeinflusst Data Governance den Wert, das Verhalten und die Governance Ihres Geschäfts?
h. Wie würden Sie sich den Zielzustand vorstellen, den CDOs oder data leaders durch die Umsetzung der Einführung von Data Governance in modernen daten-getriebenen Organisationen herbeiführen?

Welche abschließenden Gedanken möchten Sie mir noch mitgeben?

Vielen Dank für Ihre Zeit und das Gespräch.



Content of the DMBOK. Source: E. Gelbstein. Reprinted with permission.

6.4 Blank online questionnaire

The subsequent pages finally offer a full insight into the blank version of the online questionnaire which was used in order to strengthen the assumptions resulting from the preceding interviews.

Data Governance Survey

Dear participant of my survey!
My name is Christian Bruck and I am studying business informatics and business administration at the Vienna University of Economics and Business. Upon the finalization of my bachelor thesis "Challenges and opportunities of Data Governance in companies and public administration", please kindly complete the following questionnaire. Thereby you contribute significantly to the quality of the findings from my research. Answering the questions carefully takes you about 15 to 20 minutes. All data are treated anonymously and will be only used for this thesis.
Thank you very much for your support.

1. **Please indicate the type of your organization**
   Mark only one oval.

   - ◯ Large company (i.e. more than 250 employees)
   - ◯ Small and medium-sized enterprise (SME)
   - ◯ Public institution or public enterprise

2. **Please select the industry of your organization**
   Mark only one oval.

   - ◯ Agriculture, forestry and fishing
   - ◯ Mining and quarrying
   - ◯ Manufacturing
   - ◯ Electricity, gas, steam and air conditioning supply
   - ◯ Water supply; sewerage, waste management and remediation activities
   - ◯ Construction
   - ◯ Wholesale and retail trade; repair of motor vehicles and motorcycles
   - ◯ Transportation and storage
   - ◯ Accommodation and food service activities
   - ◯ Information and communication
   - ◯ Financial and insurance activities
   - ◯ Real estate activities
   - ◯ Professional, scientific and technical activities
   - ◯ Administrative and support service activities
   - ◯ Public administration and defence; compulsory social security
   - ◯ Education
   - ◯ Human health and social work activities
   - ◯ Arts, entertainment and recreation
   - ◯ Other service activities
   - ◯ Other: _____

3. **My organization has already experience with Data Governance**
   Mark only one oval.

   - ◯ Yes
   - ◯ No

4. **Please choose the option that fits best to your organization**
Mark only one oval.

- ◯ We have a small pilot project concerning Data Governance running
- ◯ Data Governance projects are running with participation of different departments within our organization
- ◯ We do not have a Data Governance project yet

5. **Are you confident with the following definition of Data Governance? If not, what should be modified?**
"Data Governance is the definition of methodological, conceptual, organizational and technical rules, responsibilities, standards and procedures aligned with the organization's strategy and culture, with the aim of using data with their maximum potential within the organizational business processes."

_____

_____

_____

_____

_____

6. **Have you heard of the Data Management Body of Knowledge (DMBOK) so far?**
(The DMBOK is one of many different models for the implementation of an effective Data Governance program).
Mark only one oval.

- ◯ Yes
- ◯ No

7. **Please select the following dimensions of Data Governance you have implemented within your organization (independent from your knowledge about the DMBOK model).**
Tick all that apply.

- ☐ Data Quality
- ☐ Metadata
- ☐ Data Warehousing and Business Intelligence
- ☐ Reference and Master Data
- ☐ Data Architecture
- ☐ Data Modelling and Development
- ☐ Data Storage and Operations
- ☐ Data Security
- ☐ Data Integration and Interoperability
- ☐ Documents and Content Management

8. **When did the implementation of your Data Governance policies begin within your organization?**
Mark only one oval.

- 2016/2017
- 2014/2015
- 2012/2013
- 2010/2011
- Before 2010
- It has not been implemented so far

9. **How long will it probably take to implement Data Governance within your whole organization?**
Mark only one oval.

- Less than 3 years
- Between 3-5 years
- Between 5-10 years
- More than 10 years
- I do not know

10. **Does your company have a particular Data Governance Framework in place or planned?**
Mark only one oval.

- Yes
- No
- In the middle of implementing one
- Just starting to design and implement one
- Other: _____

11. If yes, please describe the framework in a sentence or two:

_____

_____

_____

_____

_____

12. Did you have or still have external support for the implementation of your Data Governance (e.g. from an external consultancy firm)?
Mark only one oval.

- Yes
- No

13. **Do you have an own Data Governance department?**
Mark only one oval.

- Yes
- No
- It is planned to introduce one in the near future

14. **Responsibilities and roles**

Please choose from the list the roles already established or about to be established within your organization:
Tick all that apply.

- [ ] Data Steward
- [ ] Data Owner
- [ ] Data User
- [ ] Data Governance Coordinator
- [ ] Chief Information Officer (CIO)
- [ ] Chief Data Officer (CDO)
- [ ] Data Scientist
- [ ] Data Expert
- [ ] Data Protection Officer
- [ ] Data Management Officer
- [ ] Data Controller
- [ ] Data Manager
- [ ] Data Governance Steering Committee
- [ ] Other: _____

15. **Do you need additional staff or are you training your existing employees for the establishment of Data Governance within your organization?**

Mark only one oval.

- ◯ Additional staff
- ◯ Only train existing staff
- ◯ Both

16. **Do you work across different departments to implement Data Governance?**

Mark only one oval.

- ◯ Yes
- ◯ No

17. **Meaning and value of data**

How would you classify the importance of data for your organization?
Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Not important	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very important

18. **Data security and data protection**

Until May 2018, the EU's basic Data Protection Regulation (GDPR) must be implemented within many organizations. Please indicate the importance of the GDPR for your organization:
Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Not important	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very important

19. Do you think that a Data Governance programme could support your organization to fulfill the GDPR (or further) legal regulations?
Mark only one oval.

- ◯ Yes
- ◯ No

20. **Challenges and opportunities of Data Governance**

How would you assess the following challenges arising from an implementation of Data Governance (DG), with 1 being the lowest and 10 being the greatest challenge?
Mark only one oval per row.

	1	2	3	4	5	6	7	8	9	10
Understanding of DG concepts	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Definition of all related rules, processes and policies	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Dealing with data-related challenges from the past	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Collaboration with a consulting firm	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Complexity of a DG Framework	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Top-down or bottom-up approach	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Finding a suitable maturity model	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Lack of best practices	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Requirement of information, education, training	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Implementation takes a long time	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Cost intensity	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Overall increasing amount of data	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Size of the organization	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Effort necessary to ensure data quality	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Difficulty of a cost-benefit-weighing	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Too much everyday business inhibits an effective DG policy	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Fear coming along with such a change	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Necessity for new employees	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Focusing on data security and data protection	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Estimation of the costs involved	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Can lead to a strong bureaucratization	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Heterogeneity of the business and diverse product ranges	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Implicit and fragmentary implementation inhibits an explicit overall implementation	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Seeing data as a value-creating element	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Prospective legal obligations	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Having a cultural change within the organization	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Has not been pursued well enough until now	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯

21. Can you think of one or more further challenges your organization might be facing in the context of Data Governance?

22. **Challenges and opportunities of Data Governance**

How would you assess the following opportunities arising from an implementation of Data Governance (DG), with 1 being the lowest and 10 being the greatest opportunity?
Mark only one oval per row.

	1	2	3	4	5	6	7	8	9	10
Efficiencies in business processes										
Regulatory compliance										
Better decisions based on better data quality										
Improved customer services										
Reduced transaction costs										
Reduced customer complaints										
Ensured data security										
Consitent data quality policies										
Simplified access to relevant data										
Having pilot projects										
Having Data Scientists										
Treating data as a valuable business asset										
Having various Key Performance Indicators to measure the success of DG										
Data Lakes or Big Data projects can be better implemented										
Strengthens the external confidence with the organization										
Bridging the gap between the IT and the business										
Having the right data for the right processes										
Work-supportive function instead of being a barrier										
Minimized data redundancies										
Meeting legal data-related requirements										
Gaining value from prospective data markets										
Unifying various data-related processes										
Organizational roles are distributed more clearly										
Data-related processes are covered by responsibilities										
Enabling personalized products for customers										
Enabling more efficient gain of business related information and knowledge										
Less waste of resources										

23. Can you think of one or more further opportunites for your organization in the context of Data Governance?

_____

_____

_____

_____

_____

24. **Further new concepts**

Which of the following concepts of managing data has already been established within your organization?
Tick all that apply.

- [ ] Data Lakes
- [ ] Data Spaces
- [ ] Enterprise Linked Data
- [ ] Big Data
- [ ] None of them

25. Please asses the potential relevance of these concepts for your organization (with 1 being the lowest,10 the highest score and 0 not knowing the concept):
Mark only one oval per row.

	0	1	2	3	4	5	6	7	8	9	10
Data Lakes	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Data Spaces	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Enterprise Linked Data	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯
Big Data	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯

26. **Overall consideration, maturity, strategic importance and goals**

How important is Data Governance for your IT strategy goals?
Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Not important	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very important

27. How important is Data Governance for your business strategy goals?
Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Not important	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very important

28. How would you see the implementation stage of your organization in relation to others in your business area with regard to your Data Governance?
Mark only one oval.

- ◯ Advanced stage
- ◯ Equal stage
- ◯ Lower stage

29. How would you rate the current maturity level of your organization's Data Governance strategy (on a scale of 0-10)?
Mark only one oval.

	0	1	2	3	4	5	6	7	8	9	10	
"No formal DG process exists"	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	"The full potential of our DG has been realized"

30. How many employees work approximately in your organization?

_____

31. Do you think that Data Governance is limited to a maximum number of employees?
Mark only one oval.

- ◯ Yes, I think so
- ◯ No, I do not think so

32. **Please state the country where you work**
Mark only one oval.

- ◯ Austria
- ◯ Germany
- ◯ Switzerland
- ◯ Other: _____

33. **Are there any final thoughts you would like to share with me?**

_____

_____

_____

_____

_____