NATIONAL UNIVERSITY OF IRELAND, CORK FACULTY OF BUSINESS AND LAW

DEPARTMENT OF BUSINESS INFORMATION SYSTEMS

Critical Success Factors for Data Governance: A Theory Building Approach

Ibrahim Alhassan

112220465

Thesis submitted for the degree of Doctor of Philosophy in

Business Information Systems

Supervisors: Prof. David Sammon and Dr Mary Daly

Head of Department: Prof. Joseph Feller

May 2018

Table of Contents

Table of Contents i
List of Figures iv
List of Tables vi
Acknowledgements x
Abstract xi
Chapter One 1
1. Introduction 1
1.1 Introduction to the Study 1
1.2 Rationale behind the Study and Thesis Contributions 1
1.2.1 Research Objective and Research Questions 3
1.2.2 Overview of the Main Contributions 4
1.3 Plan of the Research 7
1.3.1 Thesis Structure 9
1.3.2 Paper 1: The Literature Review 9
1.3.3 Paper 2: The Research Method 11
1.3.4 Paper 3: Case study one 14
1.3.5 Paper 4: Case Study Two 15
1.3.6 Discussion and Conclusion 16
1.4 Data Governance 17
1.5 Research Approach 21
1.5.1 Research Philosophy 21
1.5.2 Research Strategy 23
1.5.3 Case Selection 28
1.5.4 Data Gathering 31
1.5.5 Data Analysis 36
1.6 Conclusion 39
Chapter Two 40
2. Collection of Papers 40
2.1 Paper 1 41
2.1.1 Abstract 41
2.1.2 Keywords: 42
2.1.3 Introduction 43
2.1.4 Data governance background 46
2.1.5 Research approach 49
2.1.6 Data governance activities analysis 56
2.1.6.1 Findings 61
2.1.6.2 Data governance activities model 67

2.1.7 Conclusions and Research Implications 69
2.1.7.1 Implication to theory and practice 71
2.1.7.2 Limitations and future work 72
2.1.8 References 73
2.2 Paper 2 79
2.2.1 Abstract 79
2.2.2 Keywords: 80
2.2.3 Introduction 80
2.2.4 Strauss and Corbin (1990, 1998, 2008): coding techniques 81
2.2.5 Research approach 84
2.2.6 Findings 91
2.2.6.1 Initial paper classifications 91
2.2.6.2 Pursuing OAS classification 94
2.2.6.3 Operationalising OAS coding 97
2.2.7 Discussion and recommendations for novice researchers 102
2.2.8 Concluding remarks 108
2.2.9 References 111
2.3 Paper 3 118
2.3.1 Abstract 118
2.3.2 Keywords: 118
2.3.3 Introduction 118
2.3.4 Data Governance Background 120
2.3.5 Research methodology 122
2.3.5.1 Case Background 123
2.3.5.2 Data gathering 124
2.3.5.3 Data analysis 128
2.3.6 Findings 135
2.3.6.1 CSFs for data governance 135
2.3.7 Possible interconnectedness of CSFs 145
2.3.8 Concluding Remarks and Further Research 146
2.3.9 References 148
2.4 Paper 4 151
2.4.1 Abstract 151
2.4.2 Keywords: 151
2.4.3 Introduction 152
2.4.4 Background to data governance 153
2.4.5 Research methodology 155
2.4.5.1 Case background 156
2.4.5.2 Data gathering 157

2.4.5.3 Data analysis 162
2.4.5.4 Data analysis coding procedure 164
2.4.6 Findings 167
2.4.6.1 CSFs for data governance 167
2.4.6.2 Possible interconnectedness of CSFs 177
2.4.7 Concluding remarks and future research 179
2.4.8 References 182
Chapter Three 185
3. Discussion and Conclusion 185
3.1 Introduction 185
3.2 Comparison between the Case Studies 186
3.3 CSFs for Data Governance 190
3.3.1 CSF #1: Employee data competencies 190
3.3.2 CSF #2: Flexible data tools and technologies 191
3.3.3 CSF #3: Clear data processes and procedures 193
3.3.4 CSF #4: Clear, inclusive data requirements 194
3.3.5 CSF #5: Standardised easy-to-follow data policies 195
3.3.6 CSF #6: Established data roles and responsibilities 197
3.3.7 CSF #7: Focused and tangible data strategies 198
3.3.8 CSF #8: Proper data integration strategies 199
3.3.9 CSF #9: Effective data monitoring and feedback 201
3.4 Summary of the Actions Recommended for the CSFs Identified .. 202
3.5 Possible Interconnectedness of the CSFs Identified 206
3.6 Comparison with the Literature 209
3.7 Conclusions and Research Implications 213
3.7.1 Research Contributions 216
3.7.2 Implications for Theory and Practice 219
3.7.3 Limitations and Future Research 220
References 222

List of Figures

Figure 1-1 Data governance activities model..... 11
Figure 1-2 OAS coding framwork 13
Figure 1-3 Data gathering approach (Rockart,1979). 33
Figure 2-1 illustrates the three data governance activities constructs, including possible values. 61
Figure 2-2 Comparison of the total number of concepts that emerged classified into the three actions..... 64
Figure 2-3 Comparison of the total number of concepts that emerged classified into the five decision domains..... 66
Figure 2-4 Data governance activities model..... 68
Figure 2-5 Coding framework (after Strauss and Corbin, 1990, 1998, 2008). 84
Figure 2-6 Papers that used one or more of the coding techniques, yearly distribution. 93
Figure 2-7 Context of the coding techniques. 96
Figure 2-8 Example of the relationships between the paradigm model's elements. 104
Figure 2-9 The OAS Coding Framework. 106
Figure 2-10 Illustrates the three data governance activities' constructs, including possible values (Chapter 2, paper 1). 121
Figure 2-11 Data gathering approach..... 126
Figure 2-12 Open, axial, and selective coding iterative process..... 131
Figure 2-13 Paradigm model constructs. 132
Figure 2-14 Frequency count of the categories associated with each of the five decision domains..... 133
Figure 2-15 Example of the data coding procedure for the 'Employee data competencies' CSF..... 134
Figure 2-16 illustrates the three data governance activities' constructs, including possible values (Chapter 2, paper 1). 154
Figure 2-17 Data gathering approach..... 158
Figure 2-18 Open, axial, and selective coding iterative process..... 164

Figure 2-19 Paradigm model constructs. 165
Figure 2-20 Examples of the data coding procedure for the ‘Clear, inclusive data requirements’ CSF. 166
Figure 3-1 The CSFs identified, including the percentage of each CSF in each case study. 188
Figure 3-2 Causal map of the possible interconnectedness of the nine CSFs identified. 206
Figure 3-3 Categorisation example of the nine CSFs identified. 213

List of Tables

Table 1-1 Summary of the actions recommended for the CSFs identified. 6
Table 1-2 Research plan steps. 8
Table 1-3 Data governance definitions. 19
Table 1-4 List of cases considered for this research study. 29
Table 1-5 Data preparation steps for analysis. 35
Table 2-1 Decision domains for data governance (Khatri & Brown, 2010). 48
Table 2-2 Initial classification of publications. 51
Table 2-3 Definitions of the terms that are included in open coding (adapted from Corbin & Strauss, 1990, p. 61). 55
Table 2-4 Scale of the levels of the frequency count for each type of publication. 56
Table 2-5 List of publications selected. 57
Table 2-6 Terms included in coding procedures and the total number of results. 58
Table 2-7 The concepts that emerged and their categories. 60
Table 2-8 Frequency level analysis of the data governance activities mentioned in the selected publications. 62
Table 2-9 Open, axial, and selective coding definitions from Strauss and Corbin (1990). 82
Table 2-10 Eight steps taken in data collection and analysis (after Finney and Corbett, 2007). 86
Table 2-11 Total number of excluded/included papers for each journal. 88
Table 2-12 Classification of the techniques used in the 127 papers selected. 91
Table 2-13 Analysis of the data gathering techniques used in the 59 papers. 95
Table 2-14 Classifications of the 59 papers by coding framework. 99
Table 2-15 List of interviewees' positions and related section (IT or business) and interview duration. 128
Table 2-16 Open, axial, and selective coding definitions by Strauss and Corbin (1990). 129
Table 2-17 Decision domains for data governance (Khatri & Brown, 2010). 131

Table 2-18 CSFs for data governance associated with main causes and actions/interactions..... 136
Table 2-19 Possible interconnectedness of CSFs. 146
Table 2-20 List of interviewees' positions and related sector and interview duration. 160
Table 2-21 Data preparation steps for analysis. 161
Table 2-22 Open, axial, and selective coding definitions by Strauss and Corbin (1990). 162
Table 2-23 Decision domains for data governance (Khatri & Brown, 2010). 165
Table 2-24 Possible interconnectedness of CSFs. 178
Table 3-1 Summary of the actions recommended for the 'Employee data competencies' CSF..... 191
Table 3-2 Summary of the actions recommended for the 'Flexible data tools and technologies' CSF. 192
Table 3-3 Summary of the actions recommended for the 'Clear data processes and procedures' CSF. 194
Table 3-4 Summary of the actions recommended for the 'Clear, inclusive data requirements' CSF. 195
Table 3-5 Summary of the actions recommended for the 'Standardised easy-to-follow data policies' CSF. 196
Table 3-6 Summary of the actions recommended for the 'Established data roles and responsibilities' CSF. 198
Table 3-7 Summary of the actions recommended for the 'Focused and tangible data strategies' CSF. 199
Table 3-8 Summary of the actions recommended for the 'Proper data integration strategies' CSF. 201
Table 3-9 Summary of the actions recommended for the 'Effective data monitoring and feedback' CSF. 202
Table 3-10 Illustrative examples of the maturity levels for the CSF 'Clear data processes and procedures' 203
Table 3-11 Summary of the actions recommended for the CSFs identified. 205

Table 3-12 Areas of interconnection and possible interconnectedness of the nine CSFs identified..... 208
Table 3-13 Summary of the contributions of the study..... 218

This is to certify that the work I am submitting is my own and has not been submitted for another degree, either at University College Cork or elsewhere. All external references and sources are clearly acknowledged and identified within the contents. I have read and understood the regulations of University College Cork concerning plagiarism

Acknowledgements

The completion of this thesis would not have been possible without the guidance and help of a number of individuals, who, in one way or another, contributed and extended their valuable assistance in the preparation and completion of this study.

To begin with, I would like to thank my sponsor in Saudi Arabia, Saudi Electronic University (SEU), for the scholarship and for their continuous support throughout the years of my study in Ireland.

Foremost, I would like to acknowledge the tireless and prompt help of my supervisors, Prof. David Sammon and Dr Mary Daly, for their time and effort, and for the wealth of their knowledge and experience of research. I have come to appreciate and to value the wisdom of their way of working.

To the Head of Department, Prof. Joseph Feller, for his guidance and support throughout the journey. To Prof. Ciaran Murphy, for the encouragement, listening, facilitating and much advice. To Ms Liz Eggleston, for her high-quality proofreading of my work. To my friends in Cork, who shared the difficulty of the journey and were also there to offer support.

I am deeply grateful to my family, who always encourage and support me. To my father, Prof. Mohammed, who supported me during my study by the wisdom of his recommendations. To the kindest person in my life, my mother Mody, for encouragement and love and for making me who I am today. To my brothers and sisters, who have given me positive energy during this journey.

My heartfelt thanks to my lovely wife Rana, as the person who has shared every single moment of the journey and made it easy for me. To my daughter Rand, for creating happiness and making life fun, and to my new arrival, my son Mohammed, who was born during this journey and encourages me with his lovely smiles.

Abstract

Thinking about data strategically is a challenge for many organisations today. Governing data has become vital in running a business successfully. In recent years, the volume of data used within organisations has increased dramatically, playing a critical role in business operations. The implementation of data governance remains problematic for the majority of organisations. Data governance is considered to be a relatively emerging subject, and several researchers have proposed different models that help in understanding the concepts related to it. Reviewing the literature, however, reveals a lack of research into the critical success factors (CSFs) for data governance, which shows a need for further studies aimed at understanding the success factors in governing an organisation's data.

This research study aims to identify the critical success factors for data governance that enable organisations to introduce an effective data governance programme. The research follows the building theory from case studies approach by conducting two in-depth case studies in Saudi Arabia. To gather the data, a CSF approach is employed in order to conduct the interviews. The data are then analysed using open, axial, and selective coding techniques in order to inductively identify the CSFs for data governance along with the recommended actions associated with each CSF.

This study contributes to data governance research by providing nine CSFs for data governance, as well as identifying a list of recommended actions for putting the CSFs into practice. In addition, follow a rigorous inductive research approach, two theoretical models emerged: 1) a data governance activities model, which helps in better understanding the activities related to data governance that are reported in the literature; and 2) an open, axial, and selective coding framework, which helps in understanding how to use these coding techniques when analysing qualitative data.

Chapter One

1. Introduction

1.1 Introduction to the Study

This chapter presents an introduction to the research study. It aims to provide a summary of each section of this research study, as this thesis is structured as a collection of papers, with an introductory chapter and a discussion conclusion chapter. Hence, this chapter outlines the rationale behind the study (section 1.2). It then presents the research objective and research questions in section as well as a summary of the research contributions to information systems (IS) research and practice is given in section. Section 1.3 outlines the plan of this research, which includes the thesis structure, a summary of each chapter and the papers included. Section 1.4 then introduces the concept of data governance that sets the scope for the meaning of data governance as identified in this research. Section 1.5 presents the research approach, including an introduction to the research strategy, the case selection, and the data gathering and analysis techniques followed in this research. Section 1.6 brings the chapter to a conclusion.

1.2 Rationale behind the Study and Thesis Contributions

Thinking about data strategically is a challenge for many organisations today. Governing data has become vital in running a business successfully, in order for data to be treated as a valuable asset (Khatri & Brown, 2010; Otto, 2015). In recent

years, the volume of data used within organisations has increased dramatically, play a critical role in business operations (Tallon, Ramirez, & Short, 2013). However, the implementation of data governance remains a problem for the majority of organisations (Cousins, 2016).

It is argued that a lack of trust in data can lead to the wasting of up to 50% of knowledge workers' time spent "*hunting for data*" (Redman, 2013, p. 4), whereas, when "*data is trusted, it gets shared*", which can drive higher returns on data investments (Information Builders, 2014, p. 8). Hence, the question arises: How do we ensure we are building trusted data? A recent study by Holt, Ramage, Kear, and Heap (2015) indicated that 45% of the participants, who were from the global community of database and data professionals, did not have data governance policies in place. Therefore, data governance requires more attention from stakeholders.

Although data governance is considered to be a relatively recent area (Kamioka, Luo, & Tapanainen, 2016; Rasouli, Eshuis, Trienekens, Kusters, & Grefen, 2016), several researchers have proposed different data governance models (c.f. Khatri & Brown, 2010; Otto, 2011; Panian, 2010; Weber et al., 2009; Wende, 2007). These researchers have helped our understanding of the data governance subject and in shaping its boundaries. However, more theoretical studies are needed to explore how organisations can implement data governance (Cousins, 2016). Furthermore, only a limited number of papers have examined the critical success factors (CSFs) for data governance.

Research shows that the failure rate for the development and implementation of an information system project remains high (Cecez-Kecmanovic, Kautz, & Abrahall, 2014; Doherty, Ashurst, & Peppard, 2012). The rate of failure suggests the need to focus the attention of IS professionals and academics on addressing and developing a list of the critical success factors that will enable the successful development and implementation of a new IS project (Doherty, Ashurst, & Peppard, 2012). However, providing a list of CSFs is only a partial aid to success; more is required on the implementation actions required around the list of CSFs stated (King & Burgess, 2006; Remus & Wiener, 2010; Ram & Corkindale, 2014). In terms of data governance, successful implementation can lead to the optimisation of data usage (Rifaie, Alhajj, & Ridley, 2009), which, as a consequence, delivers reliable and usable business information (CDI Institute, 2006) that enables better decision making. Having a successful data governance programme in place also supports organisations' efforts to survive by ensuring they are compliant with regulatory frameworks and able to safeguard data, particularly when a company handles sensitive material (Russom, 2008; Otto, 2011c).

Therefore, this research study aims to contribute to the IS community by filling the above-mentioned gap and identifying the CSFs for data governance, while also identifying actions that are recommended for the CSFs to be successfully put into practice. The next section outlines the research objective and research questions.

1.2.1 Research Objective and Research Questions

Given the absence of prior literature addressing the CSFs explicitly for data governance, as well as the activities included in data governance, the objective of

this research study is *"to identify the critical success factors for data governance that enable organisations to introduce an effective data governance programme"*. In order to address this research objective, two research questions have been formulated as follows:

RQ1: *What are the CSFs for data governance?*

RQ2: *What are the recommended actions for putting the CSFs for data governance into practice?*

A comprehensive review of the literature was performed and resulted in analysis of more than 80 scientific and 76 practice-oriented publications that relate to data governance. None of them explicitly study the CSFs for data governance or the data governance activities related to them. However, these publications were used to better understand data governance-related concepts and models. Therefore, it was decided to conduct an exploratory study to answer the research questions.

1.2.2 Overview of the Main Contributions

Given that this study was conducted by following the approach of building theory from case studies (c.f. Eisenhardt, 1989) within the research study journey, this research offers different contributions to data governance and IS research and practice. However, the main contribution is that this research study inductively identifies nine CSFs for data governance by analysing two case studies. A list of the recommended actions that should be taken in order to put the CSFs identified into practice successfully is provided (see Table 1-1).

During the process of conducting this research study, two additional theoretical models emerged: 1) a data governance activities model that helps to better understand the tasks that relate to data governance that are reported in the literature (see Chapter 2, paper 1); and 2) an open, axial, and selective (OAS) coding framework, which helps in understanding how to use OAS coding techniques for analysing qualitative data (see Chapter 2, paper 2).

Table 1-1 Summary of the actions recommended for the CSFs identified.

CSF	Define	Implement	Monitor
Employee data competencies	The required skills and competencies for dealing with the data for each role and responsibility.	Education and training for employees in 'how to deal with data', as well as 'increasing awareness of the importance of data'.	Employee activities and performance when using data.
Flexible data tools and technologies	Data life cycle requirements to do for integration technical needs.	An appropriate technical architecture to meet integration and the data life cycle needs.	Data life cycle, and data integration.
Clear data processes and procedures	Data capture and retrieval processes for all organisational data.	Data capture and automated validation by embedding them into business systems.	The data flow and data use.
Clear, inclusive data requirements	Data requirement standards and structure template.	An appropriate data requirement standards template.	The communication between parties regarding data requirements specification.
Standardised easy-to-follow data policies	The data regulations, access rights and privileges.	The data policies within the business systems.	The compliance with external and internal data regulations.
Established data roles and responsibilities	The data governance position and data decision rights.	The assignment of data roles to decision areas.	The clarity of data responsibilities.
Focused and tangible data strategies	Data value and objectives.	The overall data governance model.	The value of the data and the return on investment.
Proper data integration strategies	The data integration objectives.	The data infrastructure to fulfil the data integration needs.	The value of data integration.
Effective data monitoring and feedback	The key performance indicators (KPIs) for data.	Data monitoring tools within each business system.	The data performance against the KPIs for data.

1.3 Plan of the Research

The plan for this research study was driven by the road map of the building theory from case study research strategy. Eisenhardt (1989) identifies eight steps within the road map for executing theory-building from case study research: 1) Getting started, 2) Selecting cases, 3) Crafting instruments and protocols, 4) Entering the field, 5) Analysing the data, 6) Shaping a hypothesis, 7) Enfolding the literature, and 8) Reaching closure.

Table 1-2 illustrates the plan of this research study within the road map steps of the building theory from case study approach. Table 1-2 includes the main research activities for each step. Some of the activities overlap between many of the sections. This is due to the nature of the thesis structure, which is explained in detail in the next section.

Table 1-2 Research plan steps.

Step *	Research activities	Presentation section
Getting started	Identify the research gap, objective, and research questions. Understand the main related concepts.	Introduction chapter
Selecting cases	Justify the case selection	Introduction chapter
	Case one background	Paper 3
	Case two background	Paper 4
Crafting instruments and protocols	Analysis of the CSF approach and identifying the data collection procedure.	Introduction, paper 3, paper 4
Entering the field	Data collection and preparation for the analysis	Paper 3, paper 4
Analysing the data	Understanding the use of open, axial and selective coding	Paper 2
	Analysis of case one	Paper 3
	Analysis of case two	Paper 4
Shaping a hypothesis	Cross-case analysis illustrating the final list of CSFs.	Discussion and conclusion chapter
Enfolding the literature	Exploring data governance activities from the literature.	Paper 1
	Comparing the resulting actions with the data governance activities.	Discussion and conclusion chapter
Reaching closure	Description of the CSFs identified and the recommended actions. Presenting the possible interconnectedness of the CSFs.	Discussion and conclusion chapter

*Following Eisenhardt (1989).

1.3.1 Thesis Structure

This research study is structured using a series of papers. It includes three main chapters: a chapter that introduces the overall thesis. The second chapter includes a collection of four papers, which aim to present a review of the literature, the research methodology, the first case study, and the second case study. The thesis concludes with a discussion and conclusion chapter, in which the results are presented and discussed. A summary of each chapter/paper is presented in the following.

This research study starts with the current chapter, which introduces the structure of this study as well as the study objective and research questions. The remainder of this chapter covers some introductory elements of the research background and methodology that are not fully covered within the series of papers due to limits on the number of pages in the papers. The next subsections provide a brief description of each paper, followed by an outline of the discussion and conclusion chapter.

The second chapter consists of a collection of papers that aim to provide structure to the thesis. The chapter includes four papers that present the literature review, the research methodology, the first case study, and the second case study. The following subsections provide a summary description of each paper.

1.3.2 Paper 1: The Literature Review

Title: Data governance activities: a comparison between scientific and practice-oriented literature. Accepted (with minor revision) by the *Journal of Enterprise Information Management*.

The first paper reviews the prior research on data governance programmes and aims to identify the data governance activities that are reported in the literature. This paper contributes to the IS community by filling the gap identified in the literature through a categorisation of current scientific and practice-oriented publications in the domain of data governance. This categorisation is undertaken in order to understand the activities involved in data governance and to compare scientific with practice-oriented publications in terms of the activities reported.

The above research concludes with a comparison of the data governance activities that are reported in scientific publications with those that appear in practice-oriented publications. It then presents a data governance activities model (see Figure 1-1) that consists of three constructs: 1) action, plus 2) area of governance, plus 3) decision domain. The proposed data governance activities model (see Chapter 2, paper 1) can support practitioners when organising or auditing a data governance programme by helping them understand the activities involved, as well as the priorities for each activity. Furthermore, the model can be used as a conceptual framework for future field study research on data governance activities.

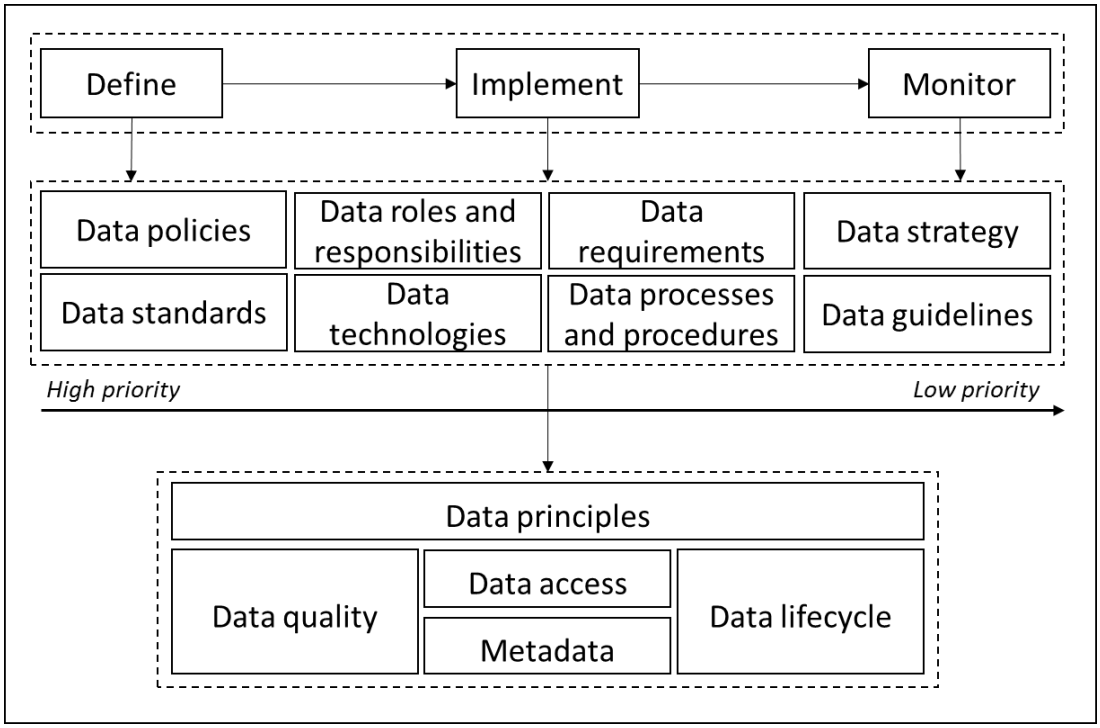

Figure 1-1 Data governance activities model.

1.3.3 Paper 2: The Research Method

Title: The use of open, axial, and selective coding techniques in IS research: a literature analysis. Under review by the *European Journal of Information Systems.*

This paper aims to investigate the use of OAS coding and is based on reviewing and analysing IS studies that have operationalised the techniques. The motivation for conducting this research was driven by the confusion that the researcher has faced while using OAS coding techniques. Hence, it is considered as a methodology paper that fits into this research study. This paper follows the structural steps taken in content analysis in order to select, review and analyse relevant literature.

The paper is intended to contribute to the IS research community by providing recommendations that will enable novice researchers to undertake OAS coding techniques proposed by Strauss and Corbin (1990). The use of coding by IS scholars has increased in recent years. However, there has been some vagueness in describing how the OAS coding techniques are executed. The paper concludes with a coding framework (see Figure 1-2) that supports the decision-making of novice researchers pursuing OAS coding as part of their qualitative data analysis. Furthermore, a list of seven items of recommended reading are presented that clearly and explicitly explain the execution of OAS coding techniques as part of data analysis.

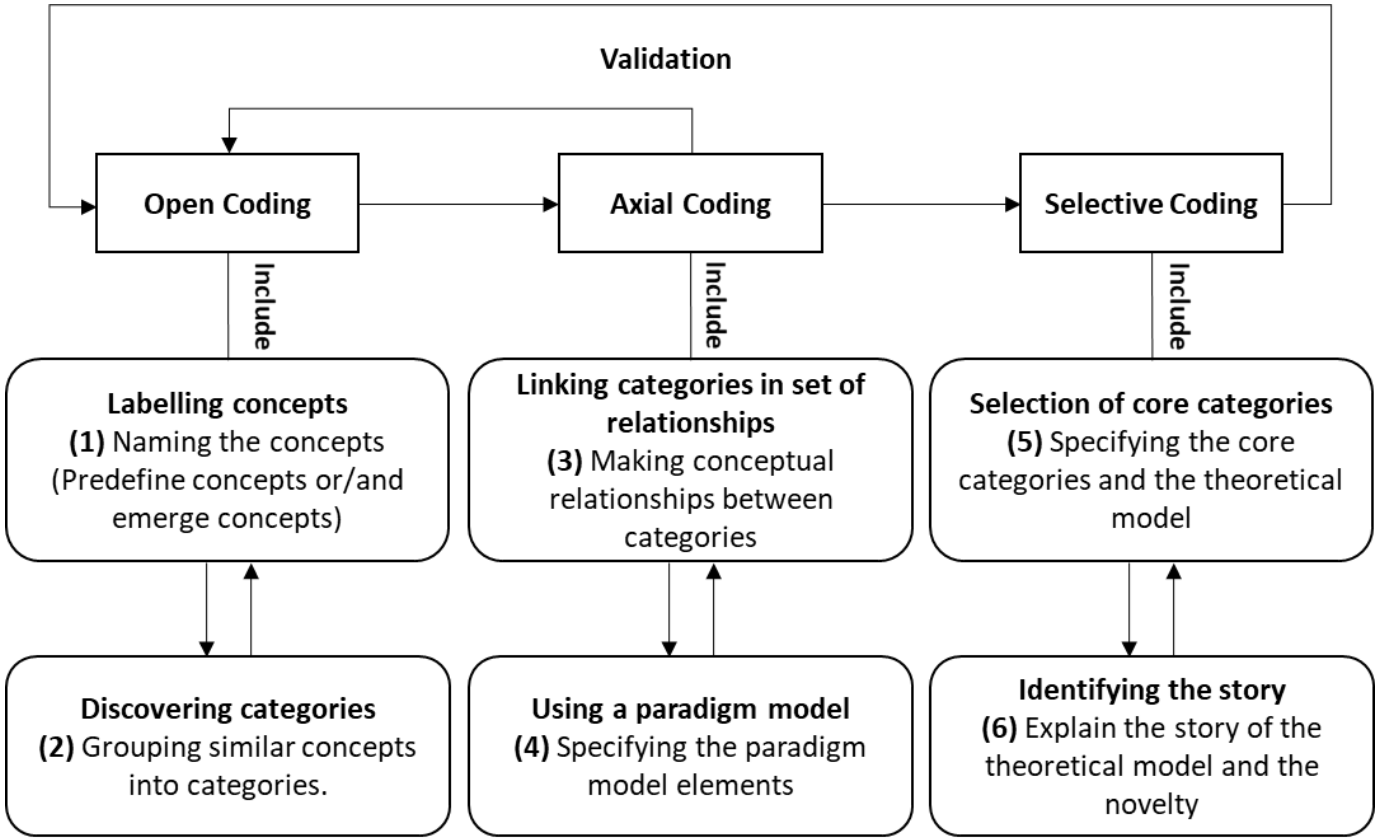

Figure 1-2 OAS coding framwork .

1.3.4 Paper 3: Case study one

Title: Critical success factors for data governance: a theory building approach. Accepted (with major revision) by *Information Systems Management*.

This paper presents the first case considered in this research study: Arajhi Bank. It aims at identifying the CSFs for data governance that emerged from analysing the first case. The data were gathered through semi-structured interviews following the CSF approach and analysed by applying open, axial, and selective coding techniques. This is considered as a within-case analysis. The findings of this research are presented as seven CSFs for data governance, which are ranked in order of importance according to the frequency count (the number of associated concepts) for each CSF. These CSFs are:

1. Employee data competencies.
2. Clear data processes and procedures.
3. Flexible data tools and technologies.
4. Standardised easy-to-follow data policies.
5. Established data roles and responsibilities.
6. Clear inclusive data requirements.
7. Focused and tangible data strategies.

This paper includes a full description of the CSFs identified from the standpoint of the first case study. The description includes the actions that are recommended to put the CSFs into practice. Finally, it highlights the relationships between the CSFs identified in order to understand their possible interconnectedness.

1.3.5 Paper 4: Case Study Two

Title: Critical success factors for data governance in the telecommunications industry. Under review for *Information and Management*.

This paper presents the second case considered in this research study: the Saudi Telecom Company (STC). It aims at identifying the CSFs for data governance that emerged from analysing the second case. The data were gathered through semi-structured interviews following the CSF approach and analysed by applying open, axial, and selective coding techniques. This is considered as a within-case analysis. The findings of this research are presented as nine CSFs for data governance, which are ranked in order of importance according to the frequency count (the number of associated concepts) for each CSF. These CSFs are:

1. Proper data integration strategies.
2. Employee data competencies.
3. Flexible data tools and technologies.
4. Clear, inclusive data requirements.
5. Clear data processes and procedures.
6. Focused and tangible data strategies.
7. Established data roles and responsibilities.
8. Accountable data access and availability.
9. Effective data monitoring and feedback.

This paper presents a full description of the CSFs identified from the standpoint of the second case. The description includes the actions that are recommended to put

the CSFs into practice. Finally, it highlights the relationships between the CSFs identified in order to understand their possible interconnectedness.

1.3.6 Discussion and Conclusion

The discussion and conclusion chapter presents a cross-case analysis and a final list of the CSFs reported by both case studies. A final list of nine CSFs is presented and each factor described. Within each CSF description, the associated recommended actions are presented and mapped to three action verbs: ‘define’, ‘implement’, and ‘monitor’.

In this chapter, the research objective of identifying the critical success factors for data governance that enable organisations to introduce an effective data governance programme is achieved by answering the two research questions: RQ1: What are the CSFs for data governance? RQ2: What are the recommended actions for putting the CSFs for data governance into practice?

The discussion section concludes with a comparison of the literature and compares the recommended actions with the data governance activities reported from Paper 1. The chapter then presents the study conclusion by summarising the findings and the contributions to theory and practice. Finally, it outlines the study limitations and presents possibilities for future work.

1.4 Data Governance

Data governance has received much attention in both the academic and practitioner communities. The concept has been developed over the last ten years, and data are now considered as valuable assets and as a strategic function within an organisation's structure (Vayghan et al., 2007; Wende, 2007). Data governance focuses on who holds the decision rights related to the data assets in an organisation (Khatri & Brown, 2010; Otto, 2011) in order to ensure the quality, consistency, usability, security, privacy, and availability of the data (Cohen, 2006; Panian, 2010).

Rau (2004, p. 35) refers to governance as "*the way the organization goes about ensuring that strategies are set, monitored, and achieved*". Horne (1995) connects governance with the optimal use of assets and outlines how data as an asset drive the importance of the governance of data within an organisation. The concept of data as an asset emerged with a report by the Hawley Committee in 1994, which defines data assets as "*data that is or should be documented and that has value or potential value*" (Oppenheim, Stenson, & Wilson, 2003, p. 159). Therefore, the main driver of data governance is the consideration of data as an asset in an organisation (Panian, 2010).

It can be argued that data governance, from both the academic and practitioner points of view, should be a universal approach to data accountability, fitting all the data aspects and needs of an organisation (Weber et al., 2009; Wende, 2007). A recent study by Holt, Ramage, Kear, and Heap (2015) indicates that 45% of their participants within the global community of database and data professionals did not

have data governance policies in place. A study by Nagle and Sammon (2017) also shows that data governance is a problem area for the majority of organisations. Hence, data governance continuously requires more attention from stakeholders (Fisher, 2006).

It can be argued that 'data governance' is a new term with novel implications for perceiving data as an asset. Several data governance models have been proposed that enable us to understand the boundaries of data governance and its related functions (Cheong & Chang, 2007; Guetat & Dakhli, 2015; Khatri & Brown, 2010; Lajara & Maçada, 2013; Otto, 2011b; Vayghan, Garfinkle, Walenta, Healy, & Valentin, 2007; Wende & Otto, 2007).

In the context of data governance, the term 'governance' should be defined and distinguished from the term 'management' in order to better understand the term 'data governance'. The main difference, therefore, between the terms 'governance' and 'management' is that governance refers to the decisions that should be made and who makes them in order to ensure the effective management and use of resources, whereas management involves implementing those decisions (Fu, Wojak, Neagu, Ridley, & Travis, 2011; Khatri & Brown, 2010). Hence, management is influenced by governance (Otto, 2011c) and, therefore, the activities for data governance can be distinguished from those required for data management.

In terms of a data governance definition, this has been presented several times in the literature (Cohen, 2006; Khatri & Brown, 2010; Loshin, 2007; Otto, 2011b; Panian, 2010; Tallon et al., 2013). However, there are differences in the definitions due to the nature of the papers' purposes in defining the term 'data governance',

Table 1-3 displays different definitions proposed in scientific and practice-oriented publications. The definitions are sorted according to the date of publication.

Table 1-3 Data governance definitions.

Reference	Definition	Publication type
Cohen (2006)	It is the process by which a company manages the quantity, consistency, usability, security and availability of data.	Academic
Loshin (2007)	It is being expected to address issues of data stewardship, ownership, compliance, privacy, data risks, data sensitivity, metadata management, master data management and even data security.	Practitioner
IBM (2007)	A quality control approach for adding new rigour and discipline to the process of managing, using, improving and protecting organisational information.	Practitioner
DAMA (2009)	The exercise of authority and control (planning, monitoring, and enforcement) over the management of data assets.	Practitioner
Panian (2010)	The processes, policies, standards, organisation, and technologies required to manage and ensure the availability, accessibility, quality, consistency, auditability, and security of data in an organisation.	Academic
Khatri and Brown (2010)	Relates to who holds the decision rights and is held accountable for an organisation's decision-making about its data assets.	Academic
Otto (2011b)	A companywide framework for assigning decision-related rights and duties in order to be able to handle data adequately as a company asset.	Academic
Tallon et al. (2013)	A collection of capabilities or practices for the creation, capture, valuation, storage, usage, control, access, archiving, and deletion of information over its life cycle.	Academic
Hall (2017)	The execution and enforcement of authority over the management of data assets and the performance of data functions.	Academic

In the light of the above definitions, it seems that data governance initiatives can be seen from different points of view. For example, Cohen (2006) narrows data governance to the processes of managing data. IBM (2007) also regards data governance as a quality control approach. With the same logic, the definition by Tallon et al. (2013) limits data governance to certain activities, rather than an overall framework for controlling data. However, decision-assigning rights is a core component of a data governance programme that supports the concept of considering data as an asset that is missing from many of the definitions.

On the other hand, Panian (2010) and Loshin (2007) generalise the data governance scope by considering different areas, such as the stewardship, processes, policies, standards, organisation, and technologies of the data. Some researchers (Khatri & Brown, 2010; Otto, 2011b; DAMA, 2009) comprehensively relate data governance to decision-holders and the authority for controlling data in an organisation. In addition, an important component and aspect of the meaning of data governance is mentioned in the definitions given by DAMA (2009), Khatri and Brown (2010) and Otto (2011b), who consider data as an organisation's asset.

Reviewing the definitions in Table 1-3, it would seem that data governance plays a fundamental role in the data and information within an organisation: it concerns the processes, policies, procedures, standards, and technology for the data aspects. It can also be noted that the purpose of a data governance programme is to manage and ensure the quality dimensions of data, such as consistency, usability, security, availability, and privacy. Such a programme also plays a role in controlling data stewardship and ownership, as well as assigning rights for whom should be acting

and making decisions with regard to any aspects of the data in an organisation. Finally, one of the most important components of data governance is considering the data in an organisation as an asset that drives value.

1.5 Research Approach

In this section, an introduction to the research approach is outlined and includes the research strategy followed. The subsections discuss and outline the case selection, data gathering, and data analysis techniques.

1.5.1 Research Philosophy

This section attempts to highlight the different views of research philosophy, starting with a discussion of the role of theory within the information systems discipline, as theory plays a fundamental role in any social science research. In the IS discipline, theory has been defined from different perspectives, summarised by Gregor (2006) as follows:

1. Theory as statements that say how something should be done in practice.
2. Theory as statements providing a lens for viewing and explaining the world.
3. Theory as statements of relationships that can be tested.

These differences in views on theory depend on philosophical and disciplinary orientations (Gregor, 2006). In the above views, theory can be seen as providing an explanation of a natural or social behaviour, event, or phenomenon (Bhattacherjee, 2012). Generally, theory is about connecting phenomena with each other,

generating a story about why acts, events, structures, and thoughts occur (Sutton & Staw, 1995).

In IS studies, research is frequently differentiated as following positivist and interpretivist/post-positivist paradigms, or using qualitative and/or quantitative methods (Gregor, 2006). The qualitative method is usually associated with interpretivism or post-positivism, whereas the quantitative method is more positivist; researchers usually use either form (Bryman, 1984). The interpretivist or post-positivist paradigm starts from the position that knowledge of reality, including the domain of human action, is a social construction by human actors (Walsham, 2006). Therefore, a post-positivist paradigm makes reasonable interpretations about a phenomenon by combining empirical data and observations of logical events, which leads to a better understanding of a social reality (Bhattacherjee, 2012).

On the other hand, the positivist paradigm treats social observations as entities in much the same way that physical scientists treat physical phenomena (Johnson & Onwuegbuzie, 2004). It has also been said that the positivist paradigm leads to a blind faith in observed data and a rejection of any attempt to extend or reason beyond observable facts (Bhattacherjee, 2012). In the IS field, the interpretivist paradigm has grown in importance (Walsham, 2006).

In addition to the first classification of research design methods, qualitative and quantitative methods are well-known classifications in any research. Both methods can be found in IS field studies, as well as mixed-methods research being used in IS research. Qualitative research methods were developed in the social sciences to

enable researchers to study social and cultural phenomena (Myers, 1997). In contrast, quantitative methods are used when the effects of an intervention on dependent variables are statistically assessed (Kaplan & Duchon, 1988). Mixed-methods research is defined by Johnson and Onwuegbuzie (2004, p.17) as "the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language into a single study".

In this research study, it was decided to employ an explanation theory in order to analyse the data qualitatively, initially by the analysis of previous literature in a rigorous approach considered to be content analysis. This enabled the researcher to review the literature comprehensively and suggest a lens for further research, as can be seen from the data governance activities model and the contribution of the data analysis framework presented in this paper. In addition, a qualitative approach was followed for the further analysis of the empirical data that were collected following a critical success factors approach.

1.5.2 Research Strategy

This research study aims to identify the CSFs for data governance inductively by conducting in-depth case studies. A case study is a research strategy that focuses on understanding the dynamics present within single settings (Eisenhardt, 1989). Benbasat, Goldstein and Mead (1987) define case study research as a process that "*examines a phenomenon in its natural setting, employing multiple methods of data collection to gather information from one or a few entities (people, groups, or organizations)*" (p. 370).

Case study is one of the most commonly used strategies in IS research, particularly for qualitative data (Darke, Shanks, & Broadbent, 1998). Research within a case study can be undertaken for either testing theory or building theory. Case study can involve a single case or multiple case studies (Eisenhardt, 1989; Yin, 2009). In a multi-case study, the results can be combined by conducting a single 'cross-case' analysis and conclusion (Yin, 2009).

Benbasat, Goldstein, and Mead (1987) have summarised the reasons for conducting case study research, in that the case study allows researchers to study a phenomenon in a natural setting, and enables them to generate theories from practice. In addition, case study research allows researchers to answer 'how' and 'why' questions, in order to understand the nature of the processes taking place. Finally, case study research is an appropriate strategy for searching an area in which limited studies have been carried out.

In order to accomplish the objective of this research study, the researcher followed the road map of building theory from case study research proposed by Eisenhardt (1989). Building theory from case study is a research strategy, for which Eisenhardt (1989) stated a clear process for conducting research that aimed to build theories. However, several other researchers had discussed different aspects of theory-building research prior to Eisenhardt (1989), such as Glaser and Strauss (1967), when initially proposed the grounded theory research as well as the development of their theory in Strauss and Corbin (1990). In addition, Yin (2009) described in depth the design of case study research as being more appropriate for a deductive research approach. However, Eisenhardt (1989) provided a road map for the

process of conducting inductive research by creating the steps and activities involved. The main driver for building theory from case study, according to Eisenhardt (1989), is when little is known about phenomena, because building theory from case studies does not rely on previous literature or prior empirical evidence.

The road map proposed by Eisenhardt (1989) involves eight steps that enable researchers to carry out rigorous research using clear processes. The eight steps are: 1) Getting started, 2) Selecting cases, 3) Crafting instruments and protocols, 4) Entering the field, 5) Analysing data, 6) Shaping hypotheses, 7) Enfolding literature, and 8) Reaching closure.

In the first step, 'Getting started', it is important to reach an initial definition of the research objective and questions, even in broad terms, in order to avoid being overwhelmed by the volume of data (Eisenhardt, 1989). The 'Selecting cases' step is a critical aspect of building theory from case study, as the case defines the set of entities from which to draw the research sample. Selecting an appropriate case study also controls unimportant variation and helps to define the limits for generalising the findings (Eisenhardt, 1989). Step three is 'Crafting instruments and protocols' and includes identifying the data gathering techniques that are to be used in order to collect the data. Qualitative data are mostly gathered in case studies by conducting interviews as gathering techniques. Both qualitative and quantitative data can be used for conducting case study research (Eisenhardt, 1989; Darke, Shanks, & Broadbent, 1998).

Step four is ‘Entering the field’, which includes an overlap between data gathering and analysis. This is important in building theory approaches, as overlapping data analysis with data collection not only gives researchers a head start but, more importantly, allows them to take advantage of flexible data collection. Flexibility enables researchers to make adjustments during the data collection processes (Eisenhardt, 1989). Therefore, this is followed by the fifth step, ‘Analysing data’, which considers the analysis within the case study. However, in the context of data analysis, Eisenhardt (1989) states: “*In fact, there are probably as many approaches as researchers*” (p. 540). Iteratively with step five, the sixth step is ‘Shaping hypotheses’, in which the emergent constructs are compared systematically with evidence from each case in order to assess how well they fit with the data. In this way, emergent constructs are measured and verified with the data (Eisenhardt, 1989).

Step seven is ‘Enfolding literature’, whereby the constructs that emerged are compared with the literature in order to understand the similarities and differences and why differences have accrued. The eighth and final step, ‘Reaching closure’, occurs when researchers stop adding more data and reach the desired theory.

In this research study, the research objective and research questions were clearly stated before starting the data collection. However, following the building theory approach does not require the prior specification of constructs. Therefore, in this research study, reviewing the prior literature was only undertaken for the purpose of understanding the boundaries and meanings of data governance programmes.

In terms of the following step, ‘Selecting cases’, this is presented in the following subsection, in which the justification for the case selection procedure is given. The third and fourth steps, ‘Crafting instruments and protocols’ and ‘Entering the field’, are then driven by the critical success factors approach (c.f., Rockart, 1979), which enables researchers to collect rigorous data that will help identify the CSFs for a phenomenon from a case study. More about the CSF approach is given in section 1.5.3 (data gathering).

Step five is ‘Analysing data’. This research study uses open, axial, and selective (OAS) coding to analyse the qualitative data gathered (c.f. Strauss & Corbin, 1990). This approach enables phenomena to emerge by following the OAS coding technique. See Chapter 2, paper 2, in which the researcher investigates the use of the OAS coding technique and concludes with the framework used to conduct the analysis of the data in this research study.

The sixth step; ‘Shaping hypotheses’, involves iterative processes that compare the emergent constructs systematically with evidence from the data. These processes are presented within the results for each case study (see Chapter 2, papers 3 and 4). The processes are also presented within the cross-case analysis (see Chapter 3), in which the final list of CSFs is presented. Following this stage, the seventh step aims at ‘Enfolding literature’ by comparing the case study result with the literature, which involves asking what is this similar to, what does it contradict, and why. In this research study, the literature review was conducted with the aim of identifying the data governance activities presented in Chapter 2, paper 1, which concludes with a data governance activities model. Finally, ‘Reaching closure’ is discussed in

the final chapter (Chapter 3), in which the final list of CSFs is presented with the actions recommended for putting the CSFs identified into practice.

1.5.3 Case Selection

Case selection is a critical decision within any case study research. The researcher should decide whether to carry out research with a single case or a multi-case study (Darke, Shanks, & Broadbent, 1998). However, multiple case selection allows for cross-case analysis and the extension of theory, as well as yielding more general research results (Benbasat, Goldstein, & Mead, 1987).

In this research study, the researcher decided to study the critical success factors for data governance in the largest industries that deal with critical data within the context of Saudi Arabia. Hence, cases were considered for inclusion in this research if they met the criteria of case selection, which were a) operating in a regulated industry, and b) having familiarity with governance practices. In general, it was found that communications and banking are the two main industries that are likely to deal with a large amount of data, and to consider data as the most critical in comparison with other sectors in Saudi Arabia. A list of cases that could be considered for inclusion in this study was compiled and is presented in Table 1-4.

Table 1-4 List of cases considered for this research study.

#	Name	Industry
1	Alrajhi Bank	Banking
2	Alahli Bank (NCB)	Banking
3	Aljazira Bank	Banking
4	Saudi Investment Bank	Banking
6	Saudi Telecom Company	Telecommunications
7	Mobily	Telecommunications

Of the seven companies nominated, it was decided to select one from the banking industry and a second case from the telecommunications sector. This enabled the researcher to conduct in-depth case study research and form a comprehensive list of CSFs for each case study.

Therefore, when comparing the list of banking industry companies, it was found that Alrajhi Bank was the most suitable case for studying CSFs for data governance. This decision was due to several factors. Mainly, Alrajhi Bank had recently approached a critical data cleansing project that involved several data governance practices. This project was aimed at complying with international regulations and requirements regarding data. Having such a project increased the awareness around data governance activities, as well as making the employees familiar with data governance practices. In addition, Alrajhi Bank is considered one of the largest banks in Saudi Arabia and has the highest number of current accounts and branches compared with other banks in the kingdom. The bank also deals with different kinds of customer relationships, which makes the customer database more complex. The

bank deals with customers as current account, investment, and remittance clients. Therefore, Alrajhi Bank seemed to be the most suitable case from the banking industry to investigate the CSFs for data governance.

For the second industry, two companies were initially considered: the Saudi Telecom Company (STC) and Mobily. It was decided to carry out the second case with STC for several reasons. Mainly, STC recently dealt with legacy data that were not governed and had attempted to integrate these with the new systems and architecture. This enabled the employees to become more familiar with the importance of data governance practices, as well as understanding the value of the data. The complexity of the data within the organisation's products and services also requires some form of data governance activities to enable the company to be effective at dealing with such complex information. STC is the largest telecommunications company in the region and serves all the cities and urban areas in Saudi Arabia. It also provides a full range of telecommunications services, which requires the management of different data infrastructures as well as mindsets. The company also recently established different services and bundles that target customers by studying their behaviour; this was enabled through a data governance programme. As part of our observations during a series of interview sessions, we could see that a data governance programme was not fully established across every area of the organisation, although a number of data governance practices were followed.

1.5.4 Data Gathering

The data gathering process in this research was inspired by the CSF approach. The CSF approach was introduced by Rockart (1979), who defines CSFs as "*areas of activity that should receive constant and careful attention from management*" (p. 85). The CSF approach has been widely investigated and used in IS research and in practice over the last three decades (Shah et al., 2007; Tan et al., 2009) and is still a valid research method for making sense of a problem by identifying potential factors that influence business (Caralli et al., 2004; Lam, 2005).

The CSF approach is a procedure that attempts to explore and identify those areas that are dictated by managerial or organisational success (Boynton & Zmud, 1984). In terms of identifying CSFs, Rockart (1979) suggests conducting two or three separate interviews with executives individually. The first round of interviews aims to identify the business goals that indicate the CSFs. Then, after analysing these goals, a set of CSFs is identified and related to the goals. The second round is used to review the CSFs identified, as well as to discuss the measures in greater depth. Finally, a third session might be required in order to obtain final agreement on the CSF measures and reporting sequence.

In addition, interviews are considered the most appropriate data gathering technique for collecting rich and detailed research material from industry experts (Koh et al., 2011). Interviews are subject to the amount of control utilised by the researcher during the interview and the degree of structure required (Esterberg, 2002).

The objective of semi-structured interviews is to explore a topic more openly and allow interviewees to express their opinions and ideas in the area being researched

(Esterberg, 2002). A semi-structured interview contains a set of key questions and the interviewer is then free to follow up with queries that relate to the studied field (Arksey & Knight, 1999). The researcher has a limited understanding or expectation of what the responses are going to be. Therefore, the ideas and areas that are covered by respondents are valuable for the researcher in this kind of data collection.

The interviews conducted for this research were aimed at identifying the business goals that indicate CSFs. Therefore, this research employed semi-structured interviews, which enabled the researcher to explore the CSFs for data governance. A data collection procedure was developed based on the CSF approach in Rockart (1979), as shown in Figure 1-3. Fifteen individual interviews were conducted in each case with personnel at the managerial levels of both the business and IT/operations departments (see Chapter 2, Papers 1 and 2 for a list of the interviewees' positions and the duration of the interviews). These interviews were conducted in two different periods. The researcher decided to stop interviewing more people at the point at which information started to be repeated and the material collected was sufficiently rich to cover the majority of the data governance aspects.

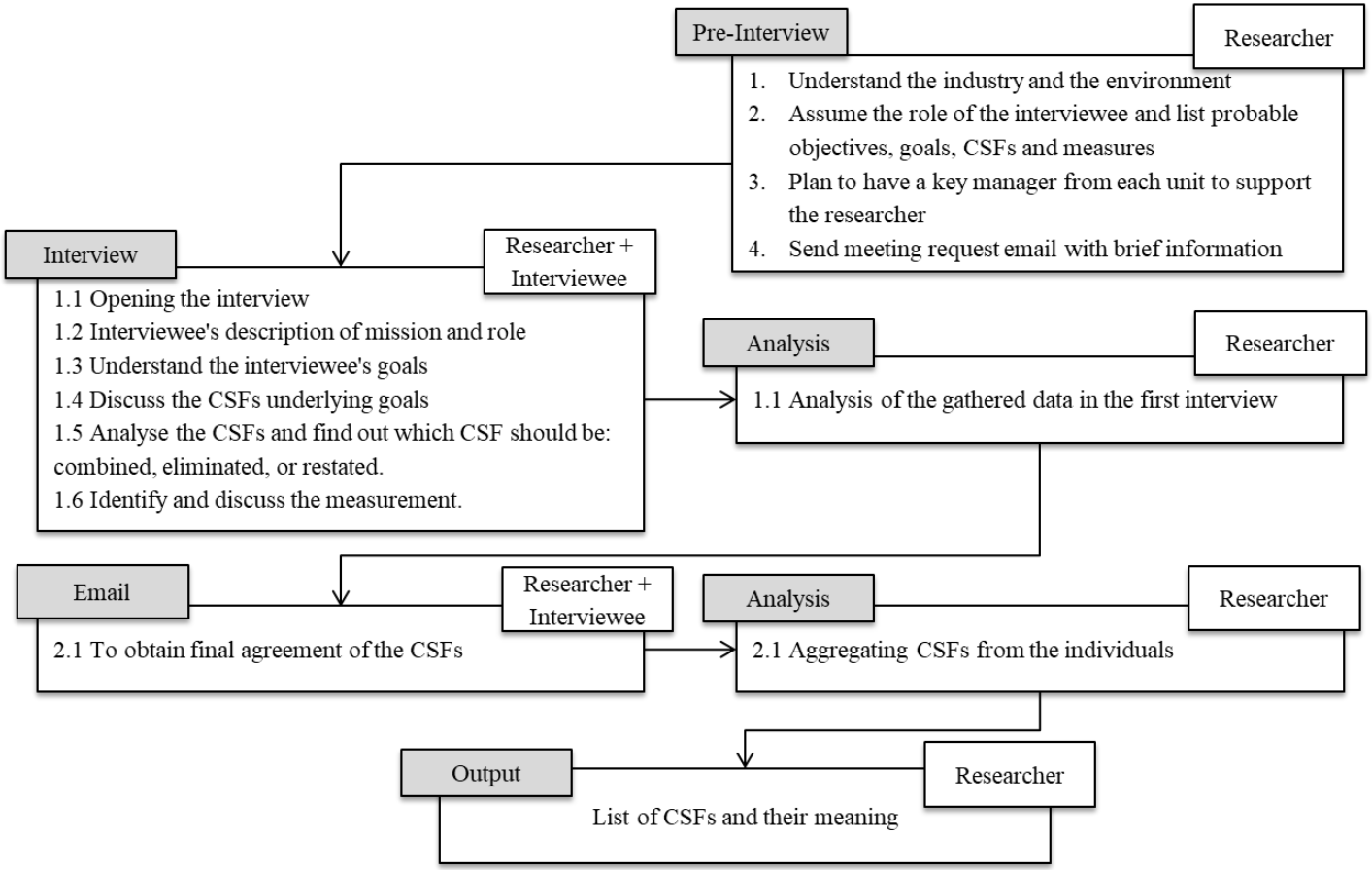

Figure 1-3 Data gathering approach (Rockart,1979).

Figure 1-3 shows six sets of activities that were followed by the researcher. The process started prior to the interviews by understanding the nature of the industry of which the case is a part as well as the case company itself. During this stage, the researcher identified the key people from each case who should be interviewed. This was done initially by searching on LinkedIn for key people at the managerial level in both the business and IT departments. Then followed by the recommendations of the interviewees.

All the interviews started with an introduction to the research objective. Each interviewee was then asked to begin talking about the data-related activities in his/her department. Then, during the interviews, we identified the CSFs related to data governance. In many cases, the interviewer explained the data governance programme from the perspective of the five decision domains to verify that the interviewee had understood the meaning of data governance. The interviewer attempted to keep the discussion to data-governance-related topics in order to concentrate the interviews around the research subject. In the second round of interviews, some of the interviewees were interviewed again in order to clarify some points from the first meeting. More participants were added in the second round based on the analysis of the first round of interviews. The individual results from each interview were also shared with them individually by email in order to be clarified.

Some of the interviews were conducted in Arabic and others in English, depending on the English-language level of the interviewee. In order to prepare the collected data for analysis, the researcher has undertaken certain steps for the data to be

analysed (see Table 1-5). All the interviews were transcribed word-by-word and those conducted in Arabic were translated into English by a third party in order to avoid bias. The transcripts were then reviewed with the recording in order to supply any missing words. Finally, due to the transcripts having been translated, they were reviewed to ensure that they were true to the meaning of the original interview.

Table 1-5 Data preparation steps for analysis.

Steps	Role	Description	Status
Record the interview	Researcher	All the interviews were recorded with the prior permission of the interviewees.	Voice only
Transcribe the interview	Third party	The interviews were listened to and the content transcribed in English. There were clear instructions written by the researcher to standardise the method for transcribing the interviews. This was done by a third party in order to avoid researcher bias.	Transcribed with 70% accuracy
Review	Researcher	The researcher listened to the interview recordings again and reviewed the transcript word-by-word to add any missing vocabulary, as well as changing or correcting phrases in order to reflect the meaning when comparing Arabic and English.	Transcribed with 100% accuracy
Data cleaning	Third party	The third party re-read the transcripts, proofread them, and reorganised them into paragraphs in order to make more sense of the data.	Ready for coding

1.5.5 Data Analysis

Data analysis may follow or overlap with data collection (Bhattacherjee, 2012), whereas following the road map referred to above (Eisenhardt, 1989) means that data collection and analysis overlap in iterative processes, as a result of which each analysis can be used for subsequent data collection in order to build a theory. It is important to consider the statement that qualitative data analysis is not well formulated (Miles, 1979) and that there are probably as many approaches as researchers (Eisenhardt, 1989). In addition, the emphasis of qualitative data analysis is on "*sense making*" (Bhattacherjee, 2012), so a coding technique was adopted in this research in a way that would serve the research objective.

Coding is one of the techniques widely used in analysing qualitative data in the IS discipline (Tallon et al., 2013). OAS coding techniques were introduced by Strauss (1987) and developed over time by Strauss and Corbin (1990, 1998, 2008) as part of the grounded theory method (Seidel & Urquhart, 2013). The coding techniques aim to generate concepts from field data (Walsham, 2006). According to Strauss and Corbin (1990, p. 57), coding "*represents the operations by which data are broken down, conceptualized, and put back together in new ways*".

Open coding is a process that aims to identify the concepts or key ideas that are hidden within data that are likely to be related to the phenomenon of interest (Bhattacherjee, 2012). Concepts and categories are generated in the open coding stage (Glaser, 1992). Strauss and Corbin (1990) state that the concepts that appear to be similar are grouped together under a higher-order, more abstract concept

called a category. When the categories are developed, their properties and the dimensions of these properties should also be identified (Strauss & Corbin, 1990).

The second phase in coding data is axial coding, which is the second reading of the data (Dezdar & Sulaiman, 2009). During this stage, the categories that emerged in the open coding stage are refined in order for them to be linked in the form of relationships. Importantly, axial coding is performed simultaneously with open coding (Strauss & Corbin, 1990; Dezdar & Sulaiman, 2009). Strauss and Corbin (1990) suggest that, in order to identify the relationships between categories, a paradigm model should be used that consists of the following elements: causal conditions, the phenomenon, the context, intervening conditions, action/interaction strategies, and consequences. Using this model enables the researcher to think systematically about the data in order to relate them (Strauss & Corbin, 1990).

Developing a paradigm model goes beyond simply developing properties and dimensions, as in the open coding stage, as the data are broken apart as concepts and become categories that have properties and dimensions; the intention is then to put the data back together in a relational form (Strauss & Corbin, 1990).

Selective coding aims to identify the core category that is the central phenomenon around which all the categories are integrated. In this stage, the analyst should be able to develop a clear storyline about the area of study (Strauss & Corbin, 1990).

Although OAS coding techniques are widely used in analysing qualitative data in the IS discipline, researchers find the use of coding both confusing and vague. Therefore, in order to evaluate the coding processes identified, the researcher conducted two different workshops that involved different levels of coding skill,

from senior lecturers at the highest level to junior PhD students at the lowest. The workshops were aimed at evaluating the coding processes. Different examples (excerpts) of the interviews were provided and the participants were asked to code each excerpt by allowing them to code all the concepts related to 'data' individually. The concepts that emerged for each excerpt were then discussed. After that, the participants started to categorise the concepts, then mapped the categories to the paradigm model in order to define the relationships and core categories.

In general, it was found that the initial results for all the participants were different in terms of labelling the concepts, as they used their own language. However, the meaning of the concepts was similar for each excerpt. In addition, there was agreement regarding the method for categorising the concepts and the paradigm model. However, there was variation in the number of concepts that emerged from each excerpt from the different participants. For example, some of the participants identified five concepts from one of the excerpts but others only found one concept in the same excerpt.

In addition to the workshops, the researcher conducted a full study that analysed the use of OAS within IS research (see Chapter 2, paper 2). Within this study, 59 papers that used OAS were analysed in order to generate the proposed framework, to map the processes for using OAS coding techniques and assist the researchers who intend to code their qualitative data using OAS coding techniques. Figure 1-2 (section 1.3.2) illustrates the OAS coding framework used within this research study.

1.6 Conclusion

This introductory chapter sets the research scope and boundaries. This chapter introduced the main elements of this research study by outlining the research objective and research questions, as well as a summary of the study contributions. This chapter also introduced the data governance concept as well as the research approach, including the research strategy, case selection, data gathering, and data analysis techniques followed.

The remainder of this thesis is structured as a collection of papers that outline the story of the research, including a review of the literature and the research methodology, in particular the data analysis; a paper that presents the findings of the first case; and a paper containing the findings of the second case. The study ends with a discussion and conclusion chapter, in which the final results are presented and discussed.

Chapter Two

2. Collection of Papers

This chapter consists of a collection of papers that aim to provide structure to the thesis. The chapter includes four papers that present the literature review (Paper 1), the research methodology (Paper 2), the first case study considered (Paper 3), and the second case study (Paper 4).

2.1 Paper 1

Data governance activities: a comparison between scientific and practice-oriented literature

2.1.1 Abstract

Purpose: This paper explores the current literature on data governance in scientific and practice-oriented publications, and is intended to provide a comparative analysis of the activities reported for data governance. Data has become a key organisational asset and data governance both a necessary and critical activity.

Design/methodology/approach: A comprehensive literature review is conducted in order to identify the published material that reflects the current state of knowledge. A systematic procedure was followed that identified 61 publications that explicitly mention data governance activities. Open coding techniques were applied to conduct content analysis, resulting in the identification of 591 concepts. A critical analysis also identified gaps in the literature.

Findings: Our analysis identified 120 data governance activities which are understood as: 'action' plus 'area of governance' plus 'decision domain' (e.g., define data policies for data quality). We define and present a data governance activities model based on our analysis. Our analysis also shows a higher volume of

data governance activities reported by practice-oriented publications that are associated with the ‘implement’ and ‘monitor’ actions of the areas of governance across the decision domains compared with scientific publications. Whereas, we found that the scientific publications focus more on defining activities. The results contribute to identifying research gaps and concerns on which ongoing and future research efforts can be focused.

Research limitations/implications: This paper is of interest to both academics and practitioners, as it helps them understand the activities associated with a data governance programme. Current literature fails to provide a comprehensive understanding of the data governance activities that are required when considering a data governance programme. Therefore, the proposed model for data governance activities can be used to give insights into these activities.

Originality/value: To the knowledge of the authors, this study is the first to explicitly consider data governance activities from both an academic and practice-oriented perspective.

Paper type: Literature review.

2.1.2 Keywords:

Data governance; data governance activities; content analysis; open coding

2.1.3 Introduction

Governance is a well-known term in business and emphasises the role of executives in representing and protecting the interests of the stakeholders (Kooper, Maes, & Lindgreen, 2011). Hence, the main role of governance is to monitor and control the behaviour of management (Kooper et al., 2011). The initial and most widely known use of the term governance within an organisation is at the corporate level, where 'corporate governance' is the set of processes, customs, policies, and laws that direct the way the business is administered or monitored to ensure that objectives are met (Porter, 2009; Kooper et al., 2011). In the field of information systems (IS), the term 'IT governance' was established (Weill & Ross, 2004) in order to apply the concept of governance to IT practices, as well as to fulfil a set of corporate governance attributes. Hence, IT governance is considered to be a subset of overall corporate governance (Wende, 2007) that focuses on information technology aspects that ensure the control and monitoring of activities in order to manage the risk that might be driven by the IT in the organisation (Kooper et al., 2011). In addition, according to Kooper et al. (2011), IT governance is partially aimed at compliance with related regulations for IT, such as Sarbanes-Oxley (USA) and Basel II (Europe).

In recent years, with the enormous increase in the use of data within organisations and those data being considered a strategic asset, the governance of data has become an initiative that should be considered by the organisation (Panian, 2010). With a similar framework to Weill and Ross (2004) on IT governance, Khatri and Brown (2010) introduced the design of a data governance framework (see Table 2-1). It

was felt that data governance and IT governance should follow principles from corporate governance (Lajara & Maçada, 2013). However, it is essential to emphasise that, in the context of governance functions in an organisation, data governance should be considered as a subset of the broader corporate governance function and aligned with IT governance (c.f. Cheong & Chang, 2007; Wende 2007; DAMA, 2009; Guetat & Dakhli, 2015).

The absence of a data governance programme may cause failure in the running of an organisation, as the worth of an organisation's data cannot be determined precisely. To know what data are worth, an organisation is required to know where the data are, how they are used, and where and when they are integrated. In recent years, the volume of data used within organisations has increased dramatically, playing a critical role in business operations (Tallon, Ramirez, & Short, 2013). In particular, data influence both operational and strategic decisions. The governance of these data has also become critical, where data are treated as a valuable asset (Khatri & Brown, 2010). Data governance has rapidly gained in popularity (Cheong & Chang, 2007; Khatri & Brown, 2010; Weber, Otto, & Österle, 2009) and is considered to be an emerging subject in the information systems (IS) field (Hagmann, 2013; Kamioka, Luo, & Tapanainen, 2016; Rasouli, Eshuis, Trienekens, Kusters, & Grefen, 2016). Practitioners also consider data governance a promising approach for enterprises to improve and maintain the quality and use of their data (Otto, 2011a).

It can be argued that data governance, from both the academic and practitioner points of view, should be a universal approach to data accountability, fitting all data

aspects and needs of an organisation (Weber et al., 2009; Wende, 2007). A survey of 200 organisations (Pierce, Dismute, & Yonke, 2008) found that 58% recognised data as a strategic asset. Whereas a recent study by Holt, Ramage, Kear, and Heap (2015) indicated that 45% of their participants within the global community of database and data professionals did not have data governance policies in place. Hence, data governance continuously requires more attention from stakeholders (Fisher, 2006).

Academics and practitioners have developed several data governance models that enable us to understand the boundaries of data governance (Cheong & Chang, 2007; Guetat & Dakhli, 2015; Khatri & Brown, 2010; Lajara & Maçada, 2013; Otto, 2011b; Vayghan, Garfinkle, Walenta, Healy, & Valentin, 2007; Wende & Otto, 2007) and part of the associated activities (DAMA International, 2009; Panian, 2010; Rifaie, Alhajj, & Ridley, 2009; Thomas, 2006; Weber et al., 2009). For example, Weber et al. (2009) proposed a contingency model for data governance and Otto (2011b) contributed a data governance organisation framework. However, none of these models mentioned explicitly consider data governance activities, although these might form part of the activities that support the proposed models. In addition, to our knowledge, few, if any publications have the activities associated with data governance with the aim of benefiting academics and practitioners in carrying out a data governance programme.

Therefore, this paper aims to contribute to the IS community by filling the gap identified in the literature through a categorisation of current scientific and practice-oriented publications in the domain of data governance. This categorisation is

undertaken in order to understand the activities involved in data governance and to compare scientific with practice-oriented publications in terms of the activities reported. These activities highlight the tasks that need to be performed in order to carry out a data governance programme. Three constructs emerged inductively, representing each of the data governance activities: 1) action, plus 2) area of governance, plus 3) decision domain. The paper concludes with a proposed data governance activities model composed of all the activities, including their order of priority.

This paper is organised as follows: section 2.1.4 presents an overview of the data governance literature and concludes with the research questions considered in this paper; section 2.1.4 describes the research approach used to conduct the literature review, including the publication selection strategy and the data analysis techniques applied; and section 2.1.6 presents the results of our analysis of the reported data governance activities from both scientific and practice-oriented publications including the data governance activities model. We conclude by addressing the limitations in this study and making recommendations for future work in the area.

2.1.4 Data governance background

Data governance is defined as ‘a companywide framework for assigning decision-related rights and duties in order to be able to adequately handle data as a company asset’ (Otto 2011b, p. 47). The main driver for data governance is considering data as an asset of the firm (Panian, 2010). Horne (1995) connected governance with optimal uses of assets, then treated data and information as an asset, which drives the importance of the governance of the data within an organisation. The concept

of data as an asset was developed in a report by the Hawley Committee in 1994, which defined data assets as ‘data that is or should be documented and that has value or potential value’ (Oppenheim, Stenson, & Wilson, 2003. p. 159).

It can be argued that ‘data governance’ is a new term with novel implications for data as an asset. However, there are many terms and approaches in the academic literature that deal with data and information under the IS field, such as total data quality management (TDQM) (Wang, 1998), data quality management (DQM) (Wang & Strong, 1996), among many different approaches and terms (Lucas, 2010; Otto, Wende, Schmidt, & Osl, 2007).

The main difference between the terms ‘governance’ and ‘management’ is that governance refers to the decisions that must be made and who makes these decisions in order to ensure effective management and use of resources, whereas management involves implementing decisions (Fu, Wojak, Neagu, Ridley, & Travis, 2011; Khatri & Brown, 2010). Hence, management is influenced by governance (Otto, 2011c). Therefore, we can distinguish between the activities for data governance and the activities required for data management.

The definition of data governance indicates who holds the decision rights and accountability regarding an enterprise’s data assets. Therefore, the decision domains should be identified in order to assign the right responsibilities and duties. In reviewing the literature relating to data governance frameworks, the framework proposed by Khatri and Brown (2010) was selected to present the decision domains that should be considered for data governance. The framework contains five interrelated decision domains: 1) data principles, 2) data quality, 3) metadata, 4)

data access, and 5) data lifecycle as shown in Table 2-1. These five decision domains follow a similar pattern to the IT governance decision domains proposed by Weill and Ross (2004). Each of the five decision domains address a set of core issues which are explained below.

Table 2-1 Decision domains for data governance (Khatri & Brown, 2010).

<table>
<tr><td colspan="3">Data principles</td></tr>
<tr><td rowspan="2">Data quality</td><td>Metadata</td><td rowspan="2">Data lifecycle</td></tr>
<tr><td>Data access</td></tr>
</table>

According to Khatri and Brown (2010), data principles are shown at the top of the framework as they are intended to establish the direction for all other decision domains. Hence, the principles set the boundary requirements for the use of data assets, which in turn addresses the enterprise's standards for data quality. The data quality then refines the basis for how data are interpreted (metadata) as well as accessed (data access) by users. Finally, the data lifecycle decision defines the production, retention and retirement of data assets which plays a fundamental role in operationalising the data principles into the IT infrastructure.

The purpose of this study is to identify and categorise the literature that explicitly mentions data governance activities in scientific publications and practice-oriented publications as well as comparing the different perspectives in order to formulate a data governance activities model. The aim is also to answer the following research questions:

RQ1: What data governance activities have been reported around the five decision domains in scientific and practice-oriented publications?

RQ2: What are the most important data governance activities presented in the scientific and practice-oriented publications?

2.1.5 Research approach

Given that the goal of this study is to gain an in-depth understanding of data governance activities, as reported in scientific and practice-oriented publications, content analysis was deemed an appropriate analysis approach. Content analysis is a frequently used technique when analysing texts (written or visual sources) especially where the meaning of the text is relatively straightforward and obvious (Myers, 2009). Content analysis requires the researcher to code the texts in a systematic way; therefore, through searching for 'structures and patterned regularities in the text' (c.f. Myers, 2009) the researcher applies a code to a unit of text that seeks to demonstrate the meaning of that text. Once coded, the resulting output can be both quantified and interpreted. Therefore, in effect, content analysis is best understood as "*a quantitative method of analysing the content of qualitative data*" (Myers, 2009, p.172). Similar to Finney and Corbett (2007) this research adopted eight coding steps in order to conduct content analysis on a selection of scientific and practice-oriented publications. These steps consist of data collection and coding procedures which enable researchers to ensure clarity and transparency

in the processes undertaken. These steps and the associated decisions are explained below.

Step 1: Decide the level of analysis

This step involves deciding what level of analysis should be conducted. The level of analysis can be a single word, a set of words, phrases, or an entire document (Finney & Corbett, 2007). The level of analysis in this research considered the entire publication in order to identify which of the publications were addressing data governance-related concepts. Therefore, the data collection phase followed a systematic literature review and was initiated by collecting publications through searches of the following databases: the AIS Electronic Library, CiteSeerX, EBSCO Online, Emerald Insight, ScienceDirect, and the ACM Digital Library. These six databases cover the majority of IS journals and conferences (Otto, 2011a). Search criteria were established in each database using the advanced search function. The keyword search criterion of having 'data governance' or 'information governance' in either the title or abstract was applied on 10th April 2017, followed by a systematic review of the references and citations of the scientific publications that resulted from the initial research. The overall data collection resulted in a total of 307 publications. The scientific publications were published in peer-reviewed academic journals and for conferences, while practice-oriented publications were published by industry associations, software vendors and analysts.

The abstracts of these publications were reviewed in order to enable the researchers to classify them in terms of scientific or practice-oriented publications, as well as to identify publications that could be excluded or included (see Table 2-2). Of the

307 publications, 151 were excluded and 156 included. The majority of the excluded publications were not related to the data governance domain. They had been published to serve a different interest of study that was not related to data governance, although they mentioned data governance in the abstract. For example, the Martin, Simons, Craven, and Betton (2014) publication, where data governance is mentioned in the context of 'there is a need for data governance in healthcare', was not considered to be directly related to the data governance domain or focusing on the study of data governance activities.

Table 2-2 Initial classification of publications.

Classification	Include/Exclude	Number of publications
Scientific publication	Include	80
Practice-oriented publications	Include	76
Not related to the data governance domain	Exclude	126
Not in English	Exclude	5
Duplicates	Exclude	20

Step 2: Decide how many concepts to code for

Here, researchers should decide whether to code text using a predefined set of concepts or develop a list of concepts incrementally during the process of coding (Finney & Corbett, 2007). For this research, the researchers decided to code concepts inductively that could be interpreted as data governance activities.

Therefore, all the concepts emerged incrementally through the processes of open coding. Each of these concepts was then categorised as a data governance activity.

Step 3: Decide whether to code for the existence or frequency of a concept

After a certain number of concepts have emerged, researchers should decide whether to code the concepts for existence or frequency (Finney & Corbett, 2007). If the concepts are coded for existence, this involves listing only the concepts that emerge. However, coding for frequency makes possible a discussion of saliency and emphasis (Finney & Corbett, 2007). For this research, it was decided to code for frequency in order to gain a deeper insight into the concepts that emerged, as well as to enable the researchers to compare the results between scientific and practice-oriented publications.

Step 4: Decide on how you will distinguish between concepts

During this step, researchers should decide whether to code the concepts exactly as they appear, or if they can be coded in some altered or collapsed form (Finney & Corbett, 2007). For this research, it was decided to follow open coding analysis techniques suggested by Corbin and Strauss (1990), whereby concepts that appear to be similar are grouped together under a higher-order, more abstract concept called a category.

Step 5: Develop rules for coding your text

It is necessary to define certain translation rules in order to ensure the consistency of the coding procedures (Finney & Corbett, 2007). The following translation rules were established and applied during the coding procedure:

- All publications were read the first time in order to code data governance activities. There should be an imperative verb that indicates that an action should be taken around data governance.
- All the concepts that emerged from the publications were compared to identify similarities and differences in order for them to be labelled together in categories.
- Once all the publications had been coded, the researchers examined the concepts that emerged and their properties within the actual text in order to ensure that they reflected the meaning of the text and that they were being related to the correct category.

Step 6: Decide what to do with 'irrelevant' information

This stage involves determining what to do with information in the text that was not coded (Finney & Corbett, 2007). Carley (1993) suggested that deleting irrelevant information can facilitate content analysis procedures by generating simplified text. In this research, the 156 publications initially included received a more in-depth review in order to identify which of the publications explicitly mention data governance activities. Of the 156 publications, only 61 explicitly mention the required or recommended data governance activities. These activities are the conditions or things that need to be performed in order to be considered as doing data governance. In order to be coded as concepts, the sentence had to contain an action (imperative verb) such as, 'define', 'establish', 'manage' and 'create'.

Step 7: Coding the text

Once the decision relating to irrelevant information is made, the coding procedure should start following the translation rules identified in step 5 (Finney & Corbett, 2007). As mentioned earlier, this research adopted an open coding analysis technique, which is part of a grounded theory approach (Corbin & Strauss, 1990). Open coding analysis is widely applied in conducting content analysis for a set of publications (Finney & Corbett, 2007; Goode & Gregor, 2009; Grahlmann, Helms, Hilhorst, Brinkkemper, & van Amerongen, 2012) and is described as 'the process of breaking down, examining, comparing, conceptualising, and categorising data' (Corbin & Strauss, 1990, p. 61). Analysing the publications using open coding enables identification of the related concepts that can be considered as emerging activities for data governance within the text of each publication within a recognised procedure.

Open coding is a process that aims to identify the concepts or key ideas that may be hidden within data and are likely to be related to a phenomenon of interest (Bhattacherjee, 2012). Concepts and categories are generated in the open coding stage (Glaser, 1992) and, when the categories are developed, their properties and the dimensions of the properties are identified (Corbin & Strauss, 1990). Table 2-3 shows the terms that are involved in open coding relevant to this study as defined by Corbin and Strauss (1990).

Table 2-3 Definitions of the terms that are included in open coding (adapted from Corbin & Strauss, 1990, p. 61).

Term	Definition
Concept	Conceptual labels placed on discrete happenings, events, and other instances of phenomena.
Category	A classification of concepts. This classification is revealed when concepts are compared one against another and appear to pertain to a similar phenomenon. Thus, the concepts are grouped together under a higher-order, more abstract concept called a category.
Coding	The process of analysing data.
Properties	Attributes or characteristics pertaining to a category.

Step 8: Analysing the results

After coding the data, researchers should decide how to review and present the results (Finney & Corbett, 2007). For this research, a frequency count was the principal method of representing the data governance activities. However, in order for the results to be compared fairly, they were translated to a scale reflecting levels of reporting namely: none, low, medium, and high. The scale levels were calculated by applying the "percentile" (Anderson, Sweeney et al. 2011) which provides information about how the data are spread over the interval from the smallest value to the largest value. The scale was for four levels for the scientific publications and other scale for the practice-oriented publications. Table 2-4 show the frequency scale for each of the levels.

Table 2-4 Scale of the levels of the frequency count for each type of publication.

	Scientific publications		Practice-oriented publications	
	From	**To**	**From**	**To**
None	0	0	0	0
Low	1	3	1	3
Medium	4	7	4	7
High	8	18	8	14

2.1.6 Data governance activities analysis

A total of 156 publications were reviewed, 61 of which explicitly mention data governance activities. The 61 publications were classified by publication type: either scientific publications, including peer-review publications, or practice-oriented publications, including publications by industry associations, software vendors and analysts. The 61 publications are listed in Table 2-5.

While reviewing and applying the open coding analysis procedure to the 61 selected publications, an MS Excel spreadsheet was developed. The spreadsheet was constructed to include a reference to each open coding stage, including referencing the original text using Mendeley, a document management application.

Table 2-5 List of publications selected.

Scientific publications (35 in total)	Al-Ruithe, Benkhelifa, & Hameed (2016); Becker (2007); Cheong & Chang (2007); Cousins, (2016); Donaldson & Walker (2004); Elliott et al., (2013); Fu et al. (2011); Gillies & Howard (2005); Guetat & Dakhli (2015); Kersten (2013); Khatri & Brown (2010); Kooper, Maes, & Lindgreen (2011); Lajara & Maçada (2013); Larkin (2008); Lomas (2010); Meyers, C. (2014); Otto (2011a, 2011b, 2011c, 2012); Palczewska et al. (2013); Panian (2010); Rickards & Ritsert (2012); Rifaie et al. (2009); Rosenbaum (2010); Shaw-Taylor (2014); Silic & Back (2013); Tallon, Ramirez, & Short (2013); Tallon, Short, & Harkins (2013); Vayghan et al (2007); Watson, Fuller, & Ariyachandra (2004); Weber et al. (2009); Weller (2008); Wende (2007); Wende & Otto (2007).
Practice-oriented publications (26 in total)	Alderson (2014); Bach (2006); Blair (2010); Bowen & Smith (2014); CDI Institute (2006); Cohen (2006); DAMA International (2009); Dember (2006); Dyché (2007); Economist Intelligence Unit (2008); Hutchinson & Sharples (2006); IBM (2007); Informatica (2013); Information Builders (2011); Khatcherian & Jefferson (2009); Loshin (2013); Moghe (2009); Nwolie (2011); Oracle (2011); Reeves & Bowen (2013); Russom (2008); Sheridan & Watzlaf (2016); Suer & Nolan (2015); The Data Warehousing Institute (2010); Thomas (2006); Wood (2013).

The open coding analysis procedure was conducted in an iterative manner (as described in step 7), starting with reading each publication and searching for any actions (imperative verbs) (see step 6). These concepts were then compared for

similarities and differences in order to categorise them into higher abstracted categories which are considered as data governance activities. Reviewing the concepts that emerged allowed us to maintain their meaning by deconstructing each of the concepts to consist of three constructs: 1) action, plus 2) area of governance, plus 3) decision domain. These constructs were later considered as the constructs of the abstracted categories (data governance activities). The actions are imperative verbs that should be undertaken within an activity, whereas the areas of governance are those aspects or functions that should have an action around them. Finally, the decision domains are one of the five referred to by Khatri and Brown (2010) in which activities are performed. Table 2-6 shows the terms used in the coding procedure associated with the number of results counted after reviewing the 61 publications. This is followed by an explanation of the three constructs.

Table 2-6 Terms included in coding procedures and the total number of results.

Term	Count	Coding example
Concept	591	Define guidelines for data quality management
Action	3	Define
Area of governance	8	Data guidelines
Decision domain	5	Data quality
Category / DG activity	120	Define data guidelines for data quality

Reviewing the 591 concepts, three ‘actions’ across the ‘areas of data governance’ emerged. These ‘actions’ indicated the doing of data governance, and were named as follows: 1) define, 2) implement, and 3) monitor. The researchers found that all the imperative verbs in the concepts could be turned into one of these three actions. For example, according to Cheong and Chang (2007, p. 1007), ‘The first step to

setting up a formal data governance programme is to determine a Data Governance structure', in which the verb 'determine' can be interpreted as 'define'. Another example comes from Weber et al. (2009, p. 4:2): 'It establishes organisation wide guidelines and standards', in which the verb 'establish' can be deemed to mean 'define'.

However, an interpretation of these actions relies upon the context itself. Therefore, each imperative verb could be interpreted as one of the three actions in one case and to another action in others, such as the verb 'develop' in some contexts means to 'define' and in others can mean 'implement'. For example, in an excerpt from Weber et al. (2009, p. 4:6), 'data governance develops and implements corporate-wide data policies', the verb 'develop' means to 'define'. In contrast, in an excerpt from Panian (2010, p. 943), 'to establish data definitions and taxonomies, define master data, develop enterprise data models', the verb 'develop' can be interpreted as 'implement', as it is related to implementing a data model.

Eight 'areas of governance' emerged inductively during the comparison procedure for the concepts that were then categorised as follows: 1) data roles and responsibilities, 2) data policies, 3) data processes and procedures, 4) data standards, 5) data strategy, 6) data technologies, 7) data guidelines, and 8) data requirements. Each of the 591 concepts could be placed into one of these areas of governance.

The third construct is the 'decision domain'. The analysis found that honouring the five decision domains defined by Kathri and Brown (2010) gave in-depth insights into the actual focus of the activity. However, some of the 591 concepts were

reported to cover more than one decision domain, and in some instances all the five decision domains. For example, a concept labelled as ‘define data policies’ without any specified domain was considered to cover all five domains.

The illustrative example below shows how the concepts were placed into a category which was considered to be a data governance activity that consisted of the three constructs. Wende (2007, p. 417) stated that ‘data governance defines roles, and it assigns responsibilities for decision areas to these roles. It establishes organisation-wide guidelines and standards for DQM’. Through coding this excerpt, four concepts emerged, which were placed into categories of data governance activities. Table 2-7 illustrates the four concepts and the breakdown of the constructs.

Table 2-7 The concepts that emerged and their categories.

Concept	**Category**		
	Action	**Area of governance**	**Decision domain**
Defines roles	Define	Data roles and responsibilities	For all decision domains
Assigns responsibilities for decision areas	Implement	Data roles and responsibilities	For all decision domains
Establishes guidelines for data quality management	Define	Data guidelines	For data quality
Establishes standards for data quality management	Define	Data standards	For data quality

During the comparison procedure, using the schema as outlined in Table 2-7, the 591 concepts were categorised into 120 data governance activities from either a

scientific or practice-oriented point of view. Figure 2-1 illustrates the three constructs with the values that emerged for each of them.

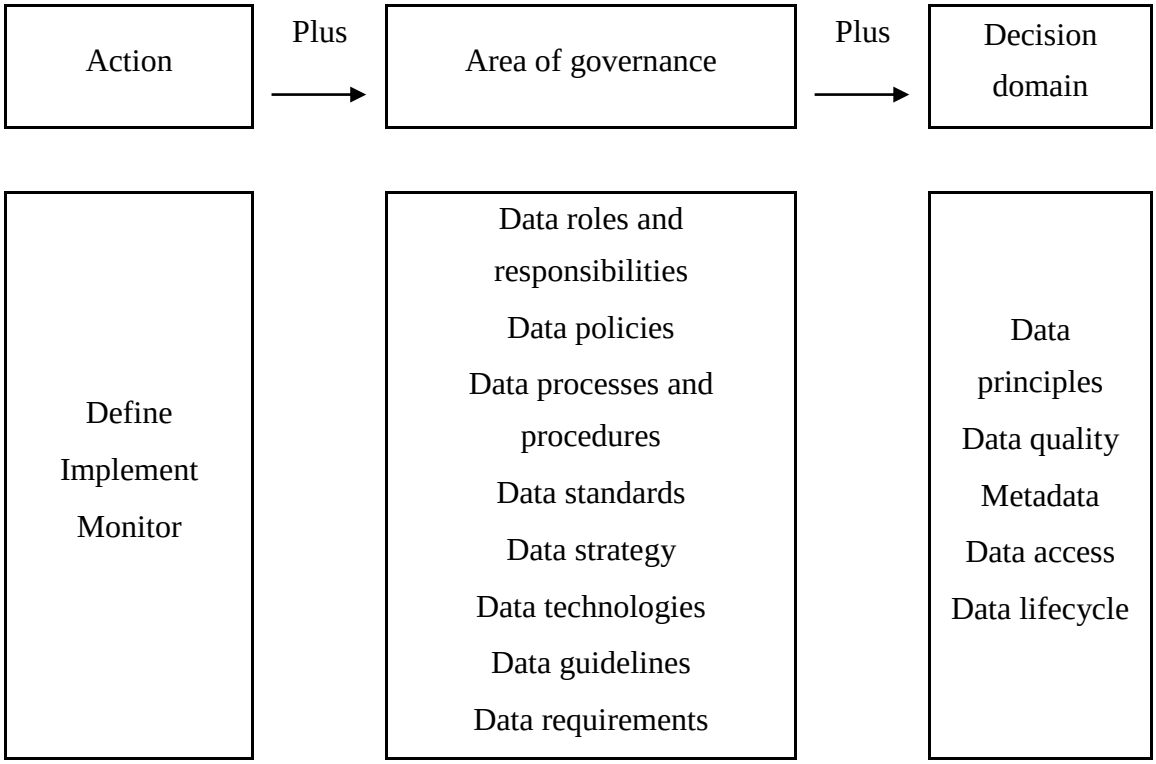

Figure 2-1 illustrates the three data governance activities constructs, including possible values.

2.1.6.1 Findings

In order to answer the first research question, Table 2-8 illustrates the results of the open coding analysis, including the level of frequency reported for the 'areas of data governance' under each respective 'action' across the five 'decision domains' for the scientific (S) publications compared with the practice-oriented (P) publications. In the event that a publication mentioned an activity more than once, the frequency was noted as '1', unless that same area was mentioned with different actions or associated with another decision domain.

Table 2-8 Frequency level analysis of the data governance activities mentioned in the selected publications.

Actions	Area of governance	Decision domains									
		Data principles		Data quality		Metadata		Data access		Data lifecycle	
		S	P	S	P	S	P	S	P	S	P
Define	Data roles and responsibilities	High	High	High	High	High	High	High	High	High	High
	Data policies	High	High	High	High	High	High	High	High	High	High
	Data processes and procedures	Low	High	High	High	Low	High	Low	High	High	High
	Data standards	High	High	High	High	High	High	High	High	High	High
	Data strategy	High	Low	High	Low	Low	Low	Low	Low	Low	Low
	Data technologies	Low	Low	Low	Low	High	Low	Low	Low	High	Low
	Data guidelines	Low	Low	Low	Medium	Low	Medium	Low	Medium	Low	Medium
	Data requirements	Low	High	Low	High	Low	High	Low	High	Low	High
Implement	Data roles and responsibilities	High	High	High	High	High	High	High	High	High	High
	Data policies	Low	High	Low	Low	Low	Low	High	High	Low	High
	Data processes and procedures	Medium	Medium	Low	Medium	Medium	Medium	Medium	Medium	Medium	Low
	Data standards	Low	Low	Low	High	Low	Low	Low	High	Low	Low
	Data strategy	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low
	Data technologies	None	Low	Medium	Low	None	Low	Medium	Low	Medium	High
	Data guidelines	Low	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Medium
	Data requirements	Medium	Low	Medium	Low	Medium	Low	Medium	Low	Low	Low
Monitor	Data roles and responsibilities	Low	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium
	Data policies	Medium	Medium	Medium	Low	Medium	Medium	Medium	Low	Medium	Low
	Data processes and procedures	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium
	Data standards	Medium	Medium	Medium	High	Medium	Low	Medium	Medium	Medium	Medium
	Data strategy	Medium	Low	Medium	Low	Medium	Medium	Medium	Medium	Medium	Medium
	Data technologies	None	Medium	None	Medium	None	Medium	None	Medium	None	Medium
	Data guidelines	Medium	None	Medium	None	Medium	None	Medium	None	Medium	None
	Data requirements	Low	High	Low	Low	Low	Low	High	High	Low	High

S: Scientific publications
P: Practice-oriented publications
* **High** * **Medium** * **Low** * **None**

Our analysis shows a considerable degree of ambiguity on the data governance activities, as none of the publications explain in detail the activities that are required to conduct a data governance programme. The actions that are reported are mostly mentioned as part of a definition of data governance or in the context of the roles and responsibilities of data governance stakeholders. Therefore, from a comprehensive view, it was found that the highest frequency count was for the area of 'data roles and responsibilities' under the 'define' and 'implement' actions across the five 'decision domains' from both scientific and practice-oriented publications. Hence, it can be argued that to 'define' and 'implement' 'data roles and responsibilities' across the five 'decision domains' is seen as the initial activity for conducting a data governance programme, as stated by Cheong and Chang (2007, p. 1007): 'The first step to setting up a formal data governance program is to determine a Data Governance structure. The structure provides escalation authority and a basis for a transparent decision making process'. The assigned roles and responsibilities will, in turn, influence how the other activities will be performed within the structure of the data governance programme..

Another noticeable finding is that the majority of the publications report activities under the 'define' action. Significantly less publications consider the 'implement' action, and only a few reported activities under the 'monitor' action. This is especially so in the scientific publications. As can be seen in Figure 2-2, the 591 concepts that emerged were classified into the three actions - 'define', 'implement' and 'monitor' - in the scientific and practice-oriented publications. There is a comparative lack of research into activities under the 'implement' and 'monitor' actions.

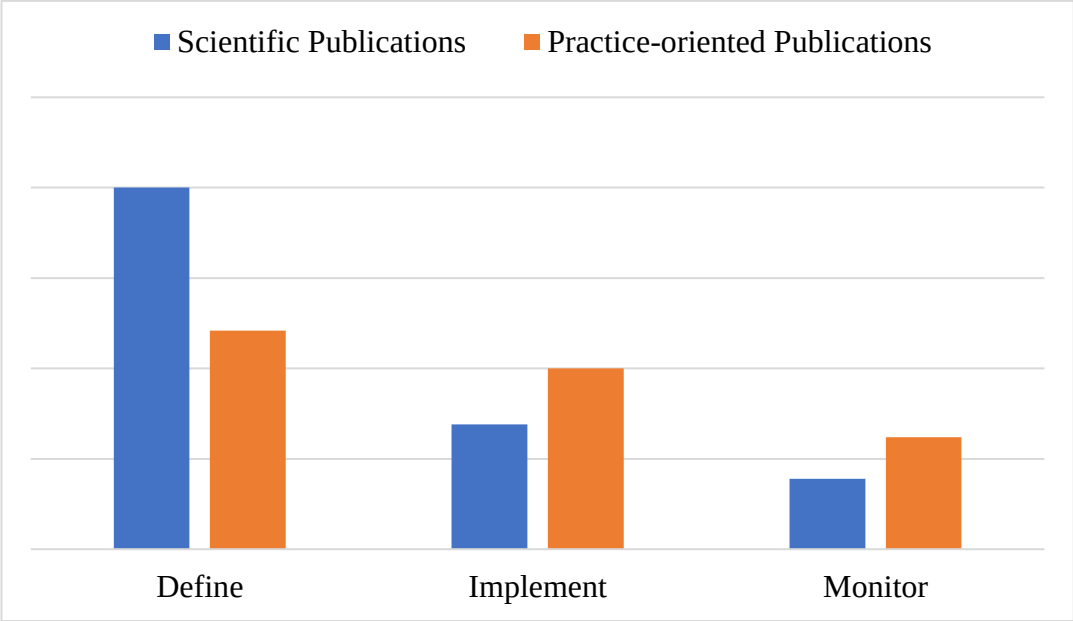

Figure 2-2 Comparison of the total number of concepts that emerged classified into the three actions.

In comparison, for the eight 'areas of governance' across the five 'decision domains', the 'define' action was reported more frequently by scientific publications than by practice-oriented publications. However, for the 'implement' and 'monitor' actions, it was observed that the practice-oriented publications focused more on these two actions compared with scientific publications. Therefore, this indicates a higher level of maturity by scientific publications in terms of defining the areas of governance across the five 'decision domains'. It can also be argued that the practice-oriented publications, particularly those from traders (such as Loshin, 2013; Russom, 2008; and Thomas, 2006), focus more on the operations aspects of a data governance programme, which are mostly under the actions of 'implement' and 'monitor'. This argument is also applicable in the case of 'data technology', as this receives more in-depth focus from practice-

oriented publications compared with scientific publications. On the other hand, ‘data requirements’ under the ‘monitor’ action receive more attention from both types of publication compared with other ‘areas of governance’. This could be due to the actual components of the ‘data requirements’, as compliance to internal and external regulations is categorised under ‘data requirements’. Therefore, because of the nature of governance, monitoring compliance with regulations is a fundamental activity for any governance type.

Additionally, as can be seen in Figure 2-2, although the total number of reported activities for the ‘define’ action is higher than for other actions in the practice-oriented publications, the difference is not as dramatic as in the scientific publications. Upon examining the 591 concepts that emerged in more detail, it was found that the majority of publications that report activities with ‘implement’ and ‘monitor’ actions had already reported the ‘define’ action in the same publication (such as DAMA International, 2009; Panian, 2010; Russom, 2008; and Wende, 2007). This comprehensive perspective provides a direction for conducting a data governance programme by focusing initially on defining the areas of governance across the five decision domains and then implementing and monitoring them.

When considering the five decision domains, many of the publications mention such activities without specifying the decision domain. For example, according to Panian (2010, p. 942), ‘It establishes the rigorous data standards’, whereby establishing data standards is considered to be the case for all the five decision domains. However, some of the reported activities explicitly focus on one of the decision domains, such as Weber et al. (2009), who mention the need to ‘develop a

corporate data quality strategy' in order to develop a data strategy for data quality as a decision domain. Figure 2-3 shows the level of focus for each of the five decision domains from scientific and practice-oriented publications.

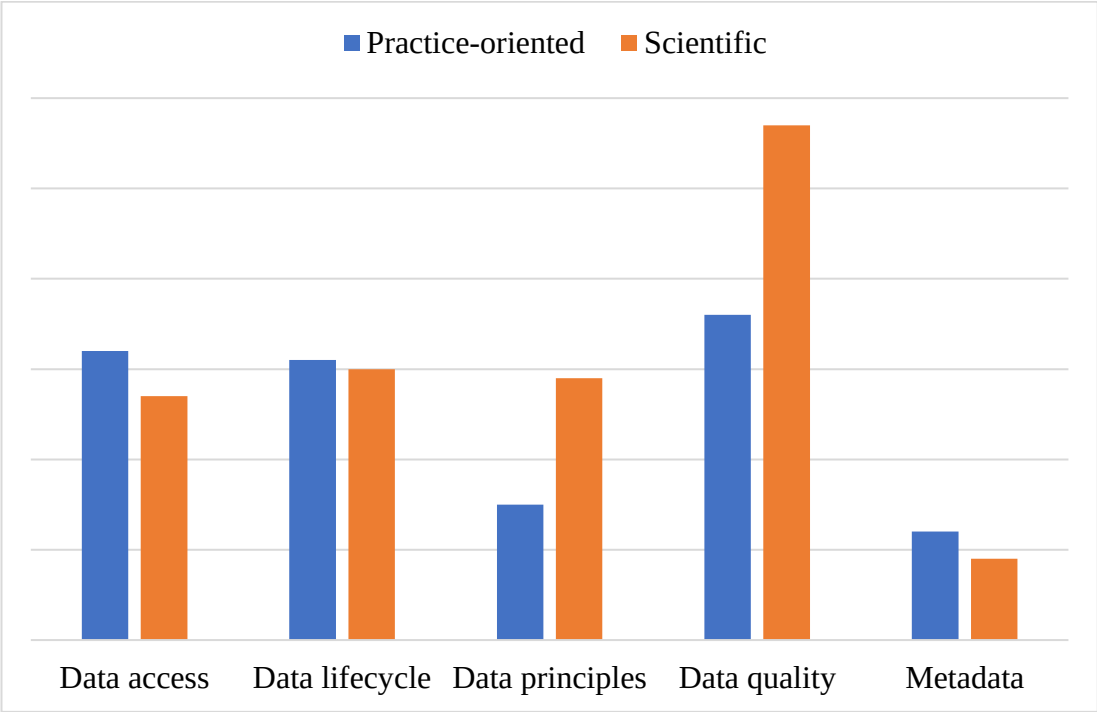

Figure 2-3 Comparison of the total number of concepts that emerged classified into the five decision domains.

It can clearly be seen in Figure 2-3 that the majority of reported activities are placed explicitly under 'data quality' as a decision domain, which is not surprising as data quality plays a fundamental role in conducting a data governance programme. It can also be argued that one of the motivations for having a data governance programme is to increase the data quality level (Otto, 2011c). However, 'data access', 'data lifecycle', and 'metadata' have been reported by practice-oriented publications more frequently than scientific publications as they are considered a technical part and more the remit of IT function. For example, Khatri and Brown

(2010, p. 149) refer to the data lifecycle as ‘Determining the definition, production, retention and retirement of data’. Therefore, the data lifecycle as a decision domain includes the technical processes (definition, production, retention, retirement of data, and more) that determine how data are treated.

2.1.6.2 Data governance activities model

The previous arguments led to the identification of data governance activities that are recommended in conducting a mature data governance programme in any organisation. This also answers the second research question: What are the most important data governance activities presented in the scientific and practice-oriented publications?

Figure 2-4 presents the data governance activities model, which consists of the three data governance activities constructs: 1) action, 2) area of governance, and 3) decision domain.

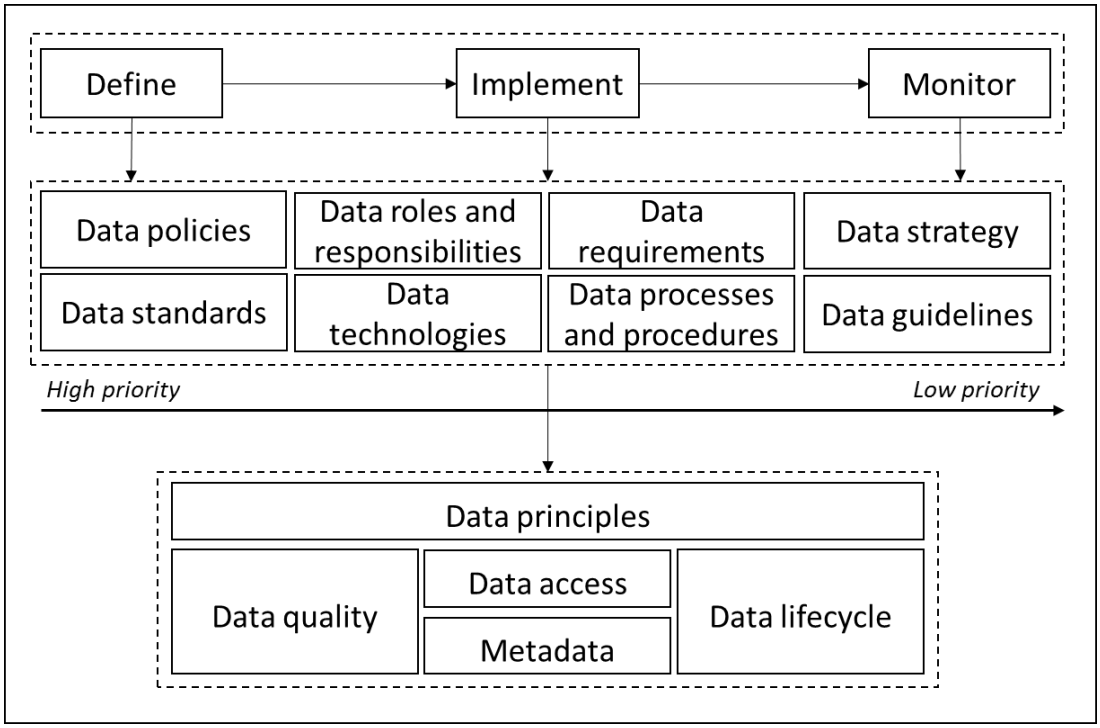

Figure 2-4 Data governance activities model.

The model recommends beginning with the activities by defining the eight areas of governance across the five decision domains. These areas of governance can then be implemented and monitored. Nevertheless, at a high level, the model suggests the priority for the areas of governance based on the frequency count from both scientific and practice-oriented publications.

The main areas of governance that overweigh the other areas from the perspective of the frequency count of the reported activities around the eight areas of governance from both scientific and practice-oriented publications are ‘data policy’, ‘data standards’, and ‘data roles and responsibilities’. These areas are reported in greater depth compared with other areas of governance. DAMA International (2009) considers data policies and data standards to be the primary

deliverables in a data governance programme. In addition, many of the publications such as (Weber et al., 2009; Wende, 2007) consider determining data roles and responsibilities as the initial activities for conducting a data governance programme. For example, the data governance model by Wende (2007) focuses only on the data roles and responsibilities in a data governance programme.

Having said that, the other areas of governance should be in place in order to conduct a data governance programme. For example, data technologies is an area of governance that receives more attention from practitioners as it is related more to the technology artefact. For example, the CDI Institute (2006, p. 12) stated that, in the context of performing data governance from an IT perspective, it involves 'developing architecture best practices and standards' and 'building governance infrastructure, technology and supporting organization' that emphasise the importance of considering the technologies that relate to conducting a data governance programme.

2.1.7 Conclusions and Research Implications

Research in the data governance domain is growing in IS, as is the need for research in this area as more organisations consider data as a valuable asset. A review of the data governance literature shows that there is a lack of research that explicitly studies activities for governing data. Nevertheless, there is some research that contributes to our understanding of data governance through modelling (Khatri & Brown, 2010; Otto, 2011b; Tallon, Ramirez, & Short, 2013). These studies reveal some progress in exploring the activities that are required for governing data.

According to Rowe (2014) there is a need within the IS community to publish more literature reviews. He argues that "*literature reviews can be highly valuable*" and "*every researcher looks for [a literature review] when starting a research study*" (Rowe, 2014, p.242). So where the main goal of a literature review is "*to classify what has been produced by the literature*" (Rowe, 2014, p.243) we believe that we have achieved this for data governance activities and mapped the territory (see Table 2-8) using the defining structure provided for a data governance activity (see Figure 2-1).

Rowe (2014, p.246) suggests that "*the quality of a literature review depends on its systematicity, since systematicity implies reproducibility through documenting the search process and potentially indicates comprehensiveness*". This research study identified and analysed 61 scientific and practice oriented-publications that focus on data governance activities. Using a systematic approach, through the eight coding steps of content analysis, the selection process yielded 307 publications that were subjected to selection and exclusion criteria, which led to the exclusion of 151. Following a more in-depth review of the remaining 156 publications, 61 were found to serve the research purpose explicitly. These 61 publications were analysed using an open coding analysis technique suggested by Corbin and Strauss (1990). This technique was selected to conduct an in-depth content analysis of the data governance activities mentioned in these publications. Therefore, we believe that we have achieved the systematicity required to ensure the reproducibility of our work by others.

2.1.7.1 Implication to theory and practice

This research concluded with a comparison of the data governance activities that are reported in scientific publications with those reported by practice-oriented publications. It was found that the scientific publications focus more on defining activities, whereas the practice-orientated publications consider the implementation and monitoring of activities. Therefore, more academic research is needed around the 'implement' and 'monitor' actions in data governance. This research concluded by presenting a data governance activities model which consists of the three constructs of data governance activities: 1) action, plus 2) area of governance, plus 3) decision domain. The proposed data governance activities model (see Figure 2-4) can support practitioners when organising or auditing a data governance programme by helping them understand the activities involved as well as the priorities for each activity. Furthermore, the model can be used as a conceptual framework for future field study research on data governance activities.

One of the main contributions of this research is the defining structure provided for a data governance activity. We argue that a data governance activity is best understood as a combination of 'action' plus 'area of governance' plus 'decision domain' (see Figure 2-1). This defining structure is a step forward in helping academics and practitioners examine the realities of data governance activities. For example, defining the data policies for data quality is very different to defining the data roles and responsibilities for data quality; therefore, our analysis and resulting activities model allows for a greater depth of understanding across data governance.

2.1.7.2 Limitations and future work

This research has two key limitations. Firstly, the research presented in this paper concluded with a frequency count of data governance activities and a data governance activities model. There is no detailed description of each of the 120 activities identified in this paper due to page length limitations. Secondly, due to the nature of this research, (a literature review), the data governance activities model that emerged has not been tested and validated through empirical research. Therefore, as a recommendation for further research, we suggest that the data governance activities should be validated by conducting field studies, as well as being described in greater detail, in order to be more valuable to both academics and practitioners.

2.1.8 References

Al-Ruithe, M., Benkhelifa, E., & Hameed, K. (2016). A conceptual framework for designing data governance for cloud computing. Procedia Computer Science, 94, 160-167.

Alderson, C 2014, 'Information governance in the NHS: Caldicott 2 and beyond', *Information System*, vol. 19, no. 6, pp. 115-119.

Anderson, DR, Sweeney, DJ, & Williams, TA 2011, *Essentials of statistics for business and economics*, revised, Cengage Learning, Stamford, Connecticut.

Bach, P 2006, 'A 360-degree approach to data governance', *Financial Executive*, vol. 22, no. 6, pp. 54-55.

Becker, MY 2007, 'Information governance in NHS's NPfIT: A case for policy specification', *International Journal of Medical Informatics*, vol. 76, no. 5-6, pp. 432-437. Available from: http://0-dx.doi.org.library.ucc.ie/10.1016/j.ijmedinf.2006.09.008.

Bhattacherjee, A 2012, *Social science research: Principles, methods, and practice*s, 2nd edn, Global Text Project, Athens, Georgia. Available from: http://scholarcommons.usf.edu/oa_textbooks/3. [22 November 2011].

Blair, BT 2010, 'Governance for protecting information in the cloud', *Information Management*, vol. 44, no. 5, pp. HT1-HT4. Available from: http://crawl.prod.proquest.com.s3.amazonaws.com/fpcache/a11e60986ad8e893b3544cc652658610.pdf?AWSAccessKeyId=AKIAJF7V7KNV2KKY2NUQ&Expires=1476691882&Signature=n5rAaY1La38h%2Fv5oL%2FQaqna%2FXZA%3D.

Bowen, R & Smith, AR 2014, 'Developing an enterprisewide data strategy', *Healthcare Financial Management*, vol. 68, no. 4, pp. 86-89.

Carley, K 1993, 'Coding choices for textual analysis: A comparison of content analysis and map analysis', *Sociological Methodology*, vol. 23, no. 75-126.

CDI Institute 2006, *Corporate data governance best practices: 2006-07 scorecards for data governance in the global 5000*, Burlingame, CA: CDI Institute.

Cheong, LK & Chang, V 2007, 'The need for data governance: A case study', *Proceedings of the 18th Australasian conference on information systems*, ACIS 2007 Proceedings, pp. 999-1008.

Cohen, R 2006, *What's in a name? Data governance roles, responsibilities and results factors*, Information Management. Available from: http://www.information-management.com/news/columns/-1057220-1.html. [8 June 2006].

Corbin, JM, & Strauss, A 1990, 'Grounded theory research: Procedures, canons, and evaluative criteria', *Qualitative Sociology*, vol. 13, no. 1, pp. 3-21.

Cousins, K. (2016). Health IT legislation in the united states: Guidelines for IS researchers. Communications of the Association for Information Systems, 39(1), 17.

DAMA International 2009, *The DAMA guide to the data management body of knowledge*, Denville, New Jersey, Technics Publications.

Dember, M 2006, '7 stages for effective data governance', *Architecture & Governance Magazine*, vol. 2, no. 4, pp. 1-3.

Donaldson, A & Walker, P 2004, 'Information governance—a view from the NHS', *International Journal of Medical Informatics*, vol. 73, no. 3, pp. 281-284. Available from: http://dx.doi.org/10.1016/j.ijmedinf.2003.11.009.

Dyché, J 2007, *A data governance manifesto: Designing and deploying sustainable data governance*, San Mateo, California: Siperian, Inc.

Economist Intelligence Unit 2008, 'The future of enterprise information governance', *The Economist*. Available from: http://viewswire.eiu.com/report_dl.asp?mode=fi&fi=1643883749.PDF.

Elliott, TE, Holmes, JH, Davidson, AJ, Chance, L, Nelson, AF, & Steiner, JF 2013, 'Data warehouse governance programs in healthcare settings: A literature review and a call to action', *EGEMs (Generating Evidence & Methods to Improve Patient Outcomes)*, vol. 1, no. 1, pp. 15.

Finney, S & Corbett, M 2007, 'ERP implementation: A compilation and analysis of critical success factors', *Business Process Management Journal*, vol. 13, no. 3, pp. 329-347. Available from: http://0-dx.doi.org.library.ucc.ie/10.1108/14637150710752272.

Fisher, T 2006, 'Data monitoring: Add controls to your data governance and compliance programs', *Business Intelligence Journal*, vol. 11, pp. 51-57.

Fu, X, Wojak, A, Neagu, D, Ridley, M., & Travis, K 2011, 'Data governance in predictive toxicology: A review', *Journal of Cheminformatics*, vol. 3, no. 1, pp. 1-16.

Gillies, A & Howard, J 2005, 'An international comparison of information in adverse events', *International Journal of Health Care Quality Assurance*, vol. 18, no. 5, pp. 343-352.

Glaser, BG 1992, *Basics of grounded theory analysis: Emergence vs. forcing*, Mill Valley, California, Sociology Press.

Goode, S & Gregor, S 2009, 'Rethinking organisational size in IS research: Meaning, measurement and redevelopment', *European Journal of Information Systems*, vol. 18, no. 1, pp. 4-25.

Grahlmann, KR, Helms, RW, Hilhorst, C, Brinkkemper, S, & van Amerongen, S 2012, 'Reviewing enterprise content management: A functional framework', *European Journal of Information Systems*, vol. 21, no. 3, pp. 268-286.

Guetat, SBA & Dakhli, SBD 2015, 'The architecture facet of information governance: The case of urbanized information systems', *Procedia Computer Science*, vol. 64, pp. 1088-1098. Available from: http://0-dx.doi.org.library.ucc.ie/10.1016/j.procs.2015.08.564.

Hagmann, J. 2013, 'Information governance – beyond the buzznull', *Records Management* Journal, vol. 23, no.3, pp. 228-240.

Holt, V, Ramage, M, Kear, K, & Heap, N 2015, 'The usage of best practices and procedures in the database community', *Information Systems*, vol, 49, pp. 163-181.

Horne, NW 1995, 'Information as an asset—The board agenda', *Computer Audit Update*, vol. 1995, no. 9, pp. 5-11. Available from: http://0-dx.doi.org.library.ucc.ie/10.1016/0960-2593(95)90246-5.

Hutchinson, C & Sharples, C 2006, 'Information governance: Practical implications for record-keeping', *Nursing Standard*, vol. 20, no.36, pp. 59-64.

IBM 2007, *The IBM data governance council maturity model: Building a roadmap for effective data governance*, Somers, New York: IBM Corporation.

Informatica 2013, *Data governance implementation: Measuring the success, compliance and effectiveness of enterprise data*, Redwood, California: Informatica.

Information Builders 2011, *Seven steps to effective data governance: A white paper*, New York, New York: Information Builders.

Kamioka, T., Luo, X., & Tapanainen, T. (2016). An empirical investigation of data governance: The role of accountabilities. *The 20th Pacific Asia Conference on Information Systems (PACIS),*

Kersten, S 2013, 'Moving the needle toward a data-driven health care system: Optimizing the EHR through information governance', *Journal of Health Care Compliance*, vol. 15, no. 3, pp. 45-62.

Khatri, V & Brown, CV 2010, 'Designing data governance', *Communications of the ACM*, vol. 53, no. 1, pp. 148-152.

Khatcherian, S & Jefferson, G 2009, 'Information governance: Challenges for financial services firms operating in the information age', *Convergence*, vol. 5, no. 1, pp. 80-89.

Kooper, MN, Maes, R, & Lindgreen, EEOR 2011, 'On the governance of information: Introducing a new concept of governance to support the management of information', *International Journal of Information Management*, vol. 31, no. 3, pp. 195-200. Available from: http://dx.doi.org/10.1016/j.ijinfomgt.2010.05.009.

Lajara, TT & Maçada, ACG 2013, 'Information governance framework: The defense manufacturing case study', *Proceedings of the nineteenth Americas conference on information systems*. Available from: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1704&context=amcis2013. [15-17 August 2013].

Larkin, BS 2008, 'Increasing information integrity: Cultural impacts of changing the way we manage data', *International Journal of Organisation Theory & Behavior (PrAcademics Press)*, vol. 11, no. 4, pp. 558-578.

Lomas, E 2010, 'Information governance: Information security and access within a UK context', *Records Management Journal*, vol. 20, no. 2, pp. 182-198.

Loshin, D 2013, *Data governance for master data management and beyond: A white paper*, Cary, North Carolina: SAS Institute Inc. Available from: http://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/data-governance-for-MDM-and-beyond-105979.pdf.

Lucas, A 2010, 'Corporate data quality management: From theory to practice', *paper presented at 5th Iberian Conference on Information Systems and Technologies (CISTI)*, Santiago de Compostela, Spain. Available from: http://hdl.handle.net/10400.5/2089. Available from: http://hdl.handle.net/10400.5/2089. [16-19 June 2010].

Martin, JL, Simons, L, Craven, MP, & Betton, V 2014, 'Towards a framework for the appraisal of digital products for mental health', *International Journal of Integrated Care*, vol. 14, no. 8, pp. 1-2. Available from: http://doi.org/10.5334/ijic.1814.

Meyers, C 2014, 'How data management and governance can enable successful self-service BI', *BI Training Solutions: As Close as Your*, vol. 19, no. 4, pp. 23.

Moghe, P 2009, 'Controlling risk with a data governance framework', *Bank Accounting & Finance*, pp. 49-51.

Myers, M. (2009) Qualitative Research in Business and Management, Sage, UK.

Nwolie, IM 2011, 'Information governance – what does it mean for you?', *Practice Nurse*, vol. 41, no. 1, pp. 1-8.

Oppenheim, C, Stenson, J, & Wilson, RMS 2003, 'Studies on information as an asset I: Definitions', *Journal of Information Science*, vol. 29, no. 3, pp. 159-166.

Oracle 2011, *Enterprise information management: Best practices in data governance*, Redwood Shores, California: Oracle.

Otto, B 2011a, 'A morphology of the organisation of data governance', *paper presented at ECIS 2011 proceedings*. Available from: http://aisel.aisnet.org/ecis2011. [6 October 2011].

Otto, B 2011b, 'Organizing data governance: Findings from the telecommunications industry and consequences for large service providers', *Communications of the Association for Information Systems*, vol. 29, no. 3, pp. 45-66. Available from: http://aisel.aisnet.org/cais/vol29/iss1/3.

Otto, B 2011c, 'Data governance', *Business & Information Systems Engineering*, vol. 3, no. 4, pp. 241-244. Available from: http://dx.doi.org.library.capella.edu/10.1007/s12599-011-0162-8.

Otto, B 2012, 'Managing the business benefits of product data management: The case of Festo', *Journal of Enterprise Information Management*, vol. 25, no. 3, pp. 272-297. Available from: http://dx.doi.org/10.1108/17410391211224426.

Otto, B, Wende, K, Schmidt, A, & Osl, P 2007, 'Towards a framework for corporate data quality management', *Proceedings of the 18th Australasian conference on information systems*, ACIS 2007 Proceedings, pp. 916-926.

Palczewska, A, Fu, X, Trundle, P, Yang, L, Neagu, D, Ridley, M, & Travis, K 2013, 'Towards model governance in predictive toxicology', *International Journal of Information Management*, vol. 33, no. 3, pp. 567-582. Available from: http://dx.doi.org/10.1016/j.ijinfomgt.2013.02.005.

Panian, Z 2010, 'Some practical experiences in data governance', World Academy of Science, Engineering and Technology Management, vol. 62, pp. 939-946.

Pierce, E, Dismute, WS, & Yonke, CL 2008, Industry report: The state of information and data governance: Understanding how organisations govern their information and data assets, Baltimore, Maryland: International Association for Information and Data Quality.

Porter, B. A. (2009). The audit trinity: The key to securing corporate accountability. Managerial Auditing Journal, 24(2), 156-182.

Rasouli, M. R., Eshuis, R., Trienekens, J. J., Kusters, R. J., & Grefen, P. W. J. (2016). Information governance as a dynamic capability in service oriented business networking. Working Conference on Virtual Enterprises, 457-468.

Reeves, MG & Bowen, R 2013, 'Developing a data governance model in health care', Healthcare Financial Management, vol. 67, no. 2, pp. 82-86.

Rickards, RC & Ritsert, R 2012, 'Data governance challenges facing controllers', *International Journal of Business, Accounting, & Finance*, vol. 6, no. 1, pp. 25-42.

Rifaie, M, Alhajj, R, & Ridley, MJ 2009, 'Data governance strategy: A key issue in building enterprise data warehouse', *paper presented at the eleventh international conference on information integration and web-based applications and services*. [14-16 December 2009].

Rosenbaum, S 2010, 'Data governance and stewardship: Designing data stewardship entities and advancing data access', *Health Services Research*, vol. 45, no. 5p2, pp. 1442-1455.

Rowe, F. (2014) What literature review is not: diversity, boundaries and recommendations, European Journal of Information Systems, 23(3), pp.241-255.

Russom, P 2008, Data governance strategies: Helping your organization comply, transform, *and integrate*, Renton, Washington: TDWI.

Shaw-Taylor, Y 2014, 'Making quality improvement programs more effective', *International J Health Care QA*, vol. 27, no. 4, pp. 264-270.

Sheridan, PT & Watzlaf, V 2016, 'Adaptive leadership in information governance', *Journal of AHIMA*, vol. 87, no. 2, pp. 20-25.

Silic, M & Back, A 2013, 'Factors impacting information governance in the mobile device dual-use contextnull', *Records Management Journal*, vol. 23, no. 2, pp. 73-89.

Suer, M & Nolan, R 12 January 2015, *Using COBIT 5 to deliver information and data governance*, COBIT Focus. Available from: http://www.isaca.org/cobit/focus/pages/using-cobit-5-to-deliver-information-and-data-governance.aspx. [7 December 2016].

Tallon, PP, Ramirez, RV, & Short, JE 2013, 'The information artifact in IT governance: Toward a theory of information governance', *Journal of Management Information Systems*, vol. 30, no. 3, pp. 141-178.

Tallon, PP, Short, JE, & Harkins, MW 2013, 'The evolution of information governance at intel', *MIS Quarterly Executive*, vol. 12, no. 4, pp. 189-198.

The Data Warehousing Institute 2010, *TDWI data governance fundamentals*, TDWI, Renton, Washington.

Thomas, G 2006, *The DGI data governance framework*, The Data Governance Institute. Available from: http://www.datagovernance.com/the-dgi-framework/. [7 December 2016].

Vayghan, JA, Garfinkle, SM, Walenta, C, Healy, DC, & Valentin, Z 2007, 'The internal information transformation of IBM', *IBM Systems Journal*, vol. 46, no. 4, pp. 669-683.

Watson, HJ, Fuller, C, & Ariyachandra, T 2004, 'Data warehouse governance: Best practices at blue cross and blue shield of north carolina', *Decision Support Systems*, vol. 38, no. 3, pp. 435-450. Available from: http://dx.doi.org/10.1016/j.dss.2003.06.001.

Wang, RY 1998, 'A product perspective on total data quality management', *Communications of the ACM*, vol. 41, no. 2, pp. 58-65.

Wang, RW & Strong, DM 1996, 'Beyond accuracy: What data quality means to data consumers', *Journal of Management Information Systems*, vol. 12, no. 4, pp. 5-33.

Weber, K, Otto, B, & Österle, H 2009, 'One size does not fit all--A contingency approach to data governance', *Journal of Data and Information Quality (JDIQ)*, vol. 1, no. 1, pp. 1-27.

Weller, A 2008, 'Data governance: Supporting datacentric risk management', *Journal of Securities Operations & Custody*, vol. 1, no. 3, pp. 250-262.

Wende, K 2007, 'A model for data governance–Organising accountabilities for data quality management', *paper presented at the 18th Australasian Conference on Information Systems*, ACIS 2007 Proceedings. Available from: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1079&context=acis2007. [5-7 December 2007].

Wende, K & Otto, B 2007, 'A contingency approach to data governance'. In: *12th International Conference on Information Quality (ICIQ-07)*, pp. 163-176. Available from: https://www.alexandria.unisg.ch/213308/. [10 November 2007].

Weill, P & Ross, JW 2004, *IT governance: How top performers manage IT decision rights for superior results*, Harvard Business School Publishing, Boston, Massachusetts.

Wood, D 2013, 'Information governance', Simpl Group Ltd, Auckland.

2.2 Paper 2

The use of open, axial, and selective coding techniques in IS research: a literature analysis

2.2.1 Abstract

Qualitative data analysis plays a critical role in academic research. Open, axial, and selective (OAS) coding techniques are widely associated with qualitative data analysis in information systems (IS) research. Therefore, this paper aims to investigate the usage of OAS coding techniques and is based on reviewing and analysing 59 published IS studies that have operationalised the techniques. This research follows the structural steps taken in content analysis in order to select, review and analyse relevant literature. The research intends to contribute to the IS research community by providing recommendations that will enable novice researchers to undertake OAS coding techniques proposed by Strauss and Corbin. Our analysis shows that the use of coding by IS scholars has increased in recent years. However, we also find that there has been some vagueness in describing how the OAS coding techniques are executed. We conclude our analysis with a coding framework that supports the decision-making of novice researchers pursuing OAS coding as part of their qualitative data analysis. Furthermore, we present a list of seven items of recommended reading that clearly and explicitly explain the execution of OAS coding techniques as part of their data analysis.

2.2.2 Keywords:

Open coding, axial coding, selective coding, data analysis, grounded theory.

2.2.3 Introduction

Open, axial, and selective (OAS) coding techniques are widely used in information system (IS) research. As part of a PhD research project there was a need to analyse qualitative data using OAS coding techniques. Although there is a considerable number of publications that explain the use of these coding techniques (c.f. Webb & Mallon, 2007; Urquhart et al., 2010; Birks et al., 2013; Matavire & Brown, 2013; Seidel & Urquhart, 2013; Urquhart & Fernandez, 2013; Wolfswinkel et al., 2013), mainly in the context of the grounded theory method, there are signs that the use of these techniques is confusing for novice researchers.

'Novice researchers' in this context refers to any researchers aiming to use OAS coding techniques for the first time. In many cases, novice researchers need to establish a clear process for how their research should be conducted (Berg, 2001) and are often unsure of how to analyse their data (Heath & Cowley, 2004). As stated by Given (2008, p. 186), "*For novice researchers, data analysis may seem like the most enigmatic and daunting aspect of qualitative research*". Hence, when novice researchers decide to conduct data analysis using OAS coding techniques, they are unsure about the procedure for operationalising these techniques, as there is confusion in this area (Eaves, 2001).

This research aims to provide recommendations to help novice researchers undertake OAS coding techniques by reviewing and analysing IS studies in top

journals (Senior Scholars' Basket of Journals)[1] that have used the techniques in different contexts. The aim of reviewing and analysing the studies is to answer the following question: How have IS scholars operationalised these coding procedures? Answering this question could help in forming recommendations for conducting OAS coding techniques.

This paper is organised as follows. It starts with an explanation of OAS coding proposed by Strauss and Corbin (1990) and concludes with a summary framework of these techniques. This is followed by the research method, whereby the research steps for content analysis are described and applied. Then, our findings are presented and the operationalisation of the OAS coding techniques is then considered. The concluding remarks are presented in the final section.

2.2.4 Strauss and Corbin (1990, 1998, 2008): coding techniques

Coding is one of the techniques widely used in analysing qualitative data in the IS discipline (Tallon et al., 2013). OAS coding techniques were introduced by Strauss (1987) and developed over time by Strauss and Corbin (1990, 1998, 2008) as part of the grounded theory method (Seidel & Urquhart, 2013).

The coding techniques referred to above aim to generate concepts from field data (Walsham, 2006). According to Strauss and Corbin (1990, p. 57), coding "*represents the operations by which data are broken down, conceptualized, and put*

[1] Web link for Senior Scholars' Basket of Journals http://aisnet.org/?SeniorScholarBasket

back together in new ways". Table 2-9 illustrates the definitions of open, axial, and selective coding according to Strauss and Corbin (1990).

Table 2-9 Open, axial, and selective coding definitions from Strauss and Corbin (1990).

Coding technique	Definition
Open coding	"*The process of breaking down, examining, comparing, conceptualizing, and categorizing data*" (p. 61).
Axial coding	"*A set of procedures whereby data are put back together in new ways after open coding, by making connections between categories. This is done by utilizing a coding paradigm involving conditions, context, action/interactional strategies and consequence*" (p. 96).
Selective coding	"*The process of selecting the core category, systematically relating it to other categories, validating those relationships, and filling in categories that need further refinement and development*" (p. 116).

Open coding is a process that aims to identify the concepts or key ideas that are hidden within data that are likely to be related to the phenomenon of interest (Bhattacherjee, 2012). Concepts and categories are generated in the open coding stage (Glaser, 1992). Strauss and Corbin (1990) state that the concepts that appear to be similar are grouped together under a higher-order, more abstract concept called a category. When the categories are developed, their properties and the dimensions of these properties should also be identified (Strauss & Corbin, 1990).

The second phase in coding data is axial coding, which is the second reading of the data (Dezdar & Sulaiman, 2009). During this stage, the categories that emerged in

the open coding stage are refined in order for them to be linked in the form of relationships. Importantly, axial coding is performed simultaneously with open coding (Strauss & Corbin, 1990; Dezdar & Sulaiman, 2009). Strauss and Corbin (1990) suggest that, in order to identify the relationship between categories, a paradigm model should be used that consists of the following elements: causal conditions, the phenomenon, the context, intervening conditions, action/interaction strategies, and consequences. Using this model enables the researcher to think systematically about the data in order to relate them (Strauss & Corbin, 1990).

Developing a paradigm model goes beyond simply developing properties and dimensions, as in the open coding stage, as the data are broken apart as concepts and become categories that have properties and dimensions; the intention is then to put the data back together in a relational form (Strauss & Corbin, 1990).

Selective coding aims to identify the core category which is the central phenomenon around which all the categories are integrated. In this stage, the analyst should be able to develop a clear story line about the area of study (Strauss & Corbin, 1990). Figure 2-5 presents a presentation of coding framework that includes the coding activities within each of the coding techniques. The coding framework below shows an iterative process between the three coding techniques, which has been visualised by understanding the work of Strauss and Corbin (1990, 1998, 2008). The coding framework is used to structure the analysis of the 59 papers reviewed as part of this research study.

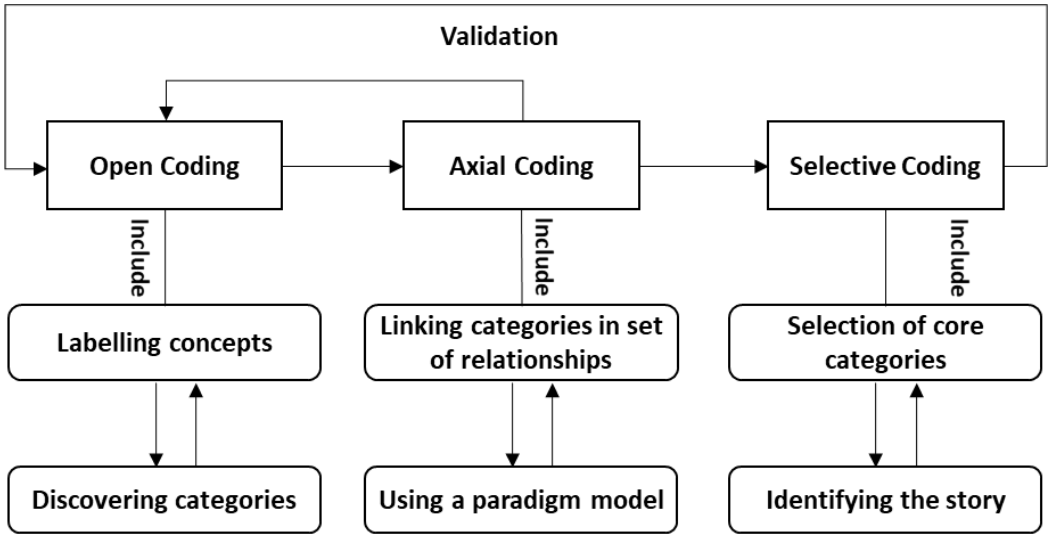

Figure 2-5 Coding framework (after Strauss and Corbin, 1990, 1998, 2008).

2.2.5 Research approach

Given that the purpose of this study is to examine how IS scholars have operationalised OAS coding techniques and to provide recommendations for novice researchers, content analysis was deemed an appropriate analysis approach. Content analysis is a frequently used technique when analysing texts (written or visual sources) especially where the meaning of the text is relatively straightforward and obvious (Myers, 2009). Content analysis requires the researcher to code the texts in a systematic way; therefore, through searching for 'structures and patterned regularities in the text' (c.f. Myers, 2009) the researcher applies a code to a unit of text that seeks to demonstrate the meaning of that text. Once coded, the resulting output can be both quantified and interpreted. Therefore, in effect, content analysis is best understood as "*a quantitative method of analysing the content of qualitative data*" (Myers, 2009, p.172). Similar to Finney and Corbett

(2007) this research adopted eight coding steps in order to conduct content analysis on a selection of scientific publications. These steps consist of data collection and coding procedures (see Table 2-10) which enable researchers to ensure clarity and transparency in the processes undertaken. These steps and the associated decisions are explained below.

Table 2-10 Eight steps taken in data collection and analysis (after Finney and Corbett, 2007).

Step	Description
Step 1: Decide the level of analysis	Researchers should decide what level of analysis should be conducted. The level of analysis can be a single word, a set of words, phrases, or an entire document.
Step 2: Decide how many concepts to code for	Researchers should decide whether to code text using a predefined set of concepts or develop a list of concepts incrementally during the process of coding
Step 3: Decide whether to code for the existence or frequency of a concept	After a certain number of concepts have emerged, researchers should decide whether to code the concepts for existence or frequency
Step 4: Decide on how you will distinguish between concepts	During this step, researchers should decide whether to code the concepts exactly as they appear, or if they can be coded in some altered or collapsed form
Step 5: Develop rules for coding your text	Researchers should define certain translation rules in order to ensure the consistency of the coding procedures
Step 6: Decide what to do with 'irrelevant' information	Researchers should determine what to do with information in the text that was not coded
Step 7: Coding the text	Researchers should start the coding procedure after the decision related to irrelevant information and should follow the translation rules identified in step 5
Step 8: Analysing the results	After coding the data, researchers should decide how to review and present the results

Step 1: Decide the level of analysis

In this research, the level of analysis was considered to be the entire research paper to identify which of the papers had either used or explained OAS coding techniques in order to be included in the initial analysis. The focus was then on the research methodology section (and specifically the data analysis section) in order to examine the use of the coding techniques.

The data collection phase was initiated by collecting papers from the Senior Scholars' Basket of eight journals. The papers were collected using the Google Scholar search engine. The Advanced function was applied for each journal. The keywords used here were ‘open coding’ OR ‘axial coding’ OR ‘selective coding’. Hence, by using these keywords as our search terms, it could be guaranteed that every paper that had applied one of the coding techniques would appear in our search. The total number of results gained from all the journals searched was 192 papers.

The data analysis section in each paper was reviewed to identify the initial relevance of the paper. The related papers in this step are those that have applied at least one of the three coding techniques. As a result, 25 papers were excluded from the total of 192. Although one or more of the techniques were mentioned in these 25 papers, this was done in an unrelated context. For example, mentioning open coding as a term but not related to the data analysis. Table 2-11 illustrates the total number of excluded and included papers for each journal.

Table 2-11 Total number of excluded/included papers for each journal.

Journal	Excluded	Included
European Journal of Information Systems	7	41
Information Systems Journal	3	28
Information Systems Research	0	10
Journal of AIS	3	15
Journal of Information Technology	2	16
Journal of MIS	1	14
Journal of Strategic Information Systems	2	21
MIS Quarterly	7	22
Total	**25**	**167**

Step 2: Decide how many concepts to code for

The papers were initially coded using the predefined concepts that are part of the coding framework (see Figure 2-5) and which were organised using a concept-centric matrix (c.f. Webster & Watson, 2002). This helped to appreciate what had been done in each paper. However, to understand how OAS coding techniques were operationalised in each paper, the researchers also decided to code concepts inductively that could simplify the use of coding analysis. Therefore, in this stage, all the concepts emerged incrementally through the processes of open coding.

Step 3: Decide whether to code for the existence or frequency of a concept

In this research, we decided to code for frequency rather than existence in order to gain a deeper insight into the concepts that emerged, as well as to avoid the uncommon use of the coding techniques.

Step 4: Decide on how you will distinguish between concepts

For this research, we decided to follow open coding analysis techniques suggested by Strauss and Corbin (1990), in which concepts that appear to be similar are grouped together under a higher-order, more abstract concept called a category.

Step 5: Develop rules for coding your text

The following translation rules were established and applied during our coding procedure:

- All papers were read the first time in order to code any relevant information about OAS coding.
- All the concepts that emerged from the papers were compared to identify similarities and differences in order for them to be labelled together in categories.
- Once all the papers had been coded, the researchers examined the categories that emerged as well as their properties within the actual text in order to ensure that they reflected the meaning of the text concerned.
- All the papers were coded within the coding framework (Figure 2-5) and organised using a concept-centric matrix.

Step 6: Decide what to do with 'irrelevant' information

The 167 papers included from step 1 underwent further review. We found that 33 of the 167 papers do not mention Strauss and Corbin (1990, 1998, 2008). Therefore, those 33 papers were considered irrelevant. The remaining 134 papers were classified into two categories: 1) those that used the coding techniques (127 papers);

and 2) those that explained the coding techniques (seven papers: Webb & Mallon, 2007; Urquhart et al., 2010; Birks et al., 2013; Matavire & Brown, 2013; Seidel & Urquhart, 2013; Urquhart & Fernandez, 2013; Wolfswinkel et al., 2013). The seven papers mainly explain the techniques within the context of the grounded theory approach. We also decided to exclude these papers as they do not use the techniques, and our focus is solely on those that have utilised the coding techniques in question. The remaining 127 papers, which used Strauss and Corbin's coding techniques, received in-depth analysis and were classified in order to direct our attention to answer our research questions of understanding how IS scholars have operationalised OAS coding techniques.

Step 7: Coding the text

We adopted the open coding technique from Strauss and Corbin (1990) for the content analysis. Open coding was used here to code any excerpt that explains the coding procedure adopted in the paper. Open coding analysis is widely applied in conducting content analysis for a set of publications (e.g. Finney & Corbett, 2007; Goode & Gregor, 2009; Grahlmann et al., 2012).

Step 8: Analysing the results

The analysis is presented under the findings and discussion sections. The findings section shows how IS scholars conduct coding techniques and includes some descriptive statistics about the papers analysed. This is followed by the method the scholars used to operationalise the three coding techniques, concluding with the recommended list of papers and coding specifications that facilitate the conduct of

the three coding techniques. These recommendations rely heavily on the frequency and similarities of how IS scholars have operationalised the coding techniques.

2.2.6 Findings

2.2.6.1 Initial paper classifications

Reviewing and analysing the initial set of papers (127 selected papers) produced interesting findings. First, while the coding techniques proposed by Strauss and Corbin are still used by IS scholars in IS studies, not all of the papers operationalised all three coding techniques. Table 2-12 shows the classification of papers in terms of their use of OAS coding techniques. It can be seen that the first classification, 'OAS', has 59 papers that have used OAS coding techniques. We focus on these 59 papers for the remainder of the analysis being reported in this paper (please see the next section).

Table 2-12 Classification of the techniques used in the 127 papers selected.

Classification	**Coding technique**			**Number of papers**
	Open	**Axial**	**Selective**	
OAS	Yes	Yes	Yes	59
OA	Yes	Yes	No	24
O	Yes	No	No	35
OS	Yes	No	Yes	4
A	No	Yes	No	4
S	No	No	Yes	1

Furthermore, Table 2-12 shows that the majority of the 127 papers are under the first three classifications in which the open coding stage is included. This is not surprising, as open coding is considered the first stage in coding qualitative data. The outcome of open coding can be considered either as the input to axial coding or, in many cases, as the final stage of the data analysis, such as in the work of Huang et al. (2003) and Benbunan-Fich and Benbunan (2007). Open coding can also be a form of initial coding for other analysis techniques. For example, Feldman and Horan (2011) conducted open coding to categorise data for the second stage of their analysis. In addition, papers under the 'OA' classification, in which no selective coding is conducted, mainly arrived at the results after axial coding was carried out (such as Olsson et al., 2008). However, some papers use the result to conduct a third stage of analysis that does not involve selective coding (such as Scott, 2000 and Furneaux and Wade, 2011).

In the 'OS' classification (axial coding missing) the four papers that applied open and selective coding without axial coding follow Glaser's (1992) coding techniques and simply mention Strauss and Corbin in order to explain open and selective coding techniques (such as Miranda et al., 2015). Glaser's approach to coding techniques excludes axial coding, as he argues that the paradigm model forces data, hinders emergence, and leads to conceptual description rather than grounded theory (Seidel & Urquhart, 2013).

However, although the 'A' and 'S' classifications, in which no open coding techniques are used, are very low volume, they do not seem to make sense initially. Taken for example the papers in which only axial coding taken place (c.f. Shang &

Seddon, 2002; Levina & Ross, 2003; Goo et al., 2009; Leonardi, 2012), the authors use the power of axial coding techniques to form relationships, using the results of the first analysis technique that was adopted. For example, one paper (Leonardi, 2012, p. 757) states, in the context of data analysis: "*In Step 3, I used the process of axial coding (Strauss and Corbin 1998) to create new codes that linked together codes from Step 1 about what informants initially heard about CrashLab to codes from Step 2 about how they used it for the first time*". Therefore, this shows that, the axial coding technique can be used to identify relationships between data that are analysed using different techniques (other than open coding).

Figure 2-6 illustrates the yearly distribution of the papers included in Table 2-12 (127 papers) with the distinction of the first classification 'OAS'. It can be seen that the overall number of papers that used the coding proposed by Strauss and Corbin has increased over time, which is an indication of the popularity and power of these techniques.

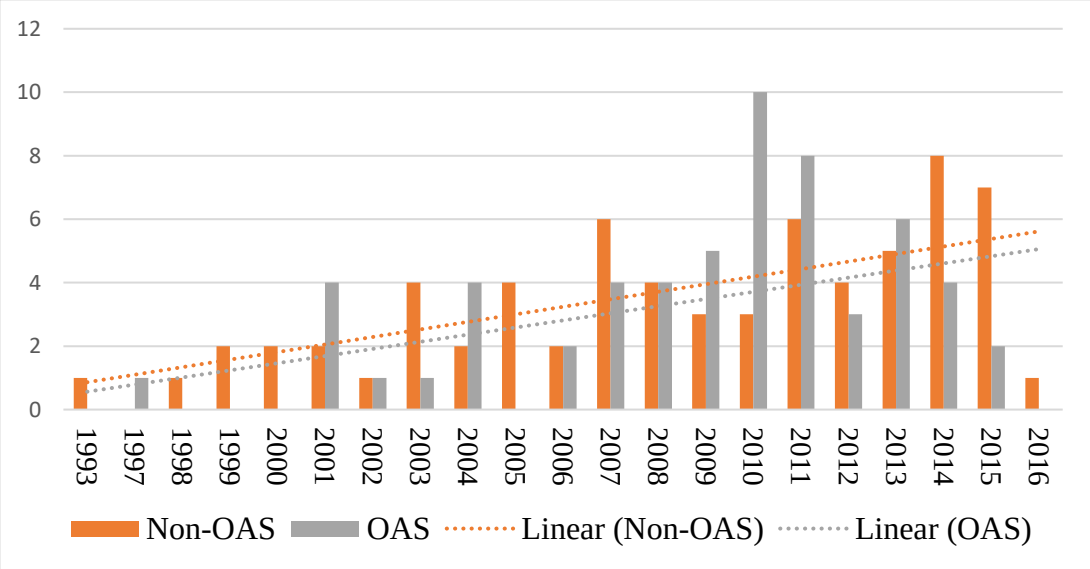

Figure 2-6 Papers that used one or more of the coding techniques, yearly distribution.

2.2.6.2 Pursuing OAS classification

The ‘OAS’ classification contains 59 papers in total (see Table 2-12). We analysed these papers using a concept-centric matrix that consists of the data gathering techniques, where they were published, as well as an analysis of the context of the coding techniques, in which there are three possible values:

1. Grounded theory (GT) approach: in which the main theory followed is grounded theory and, therefore, grounded theory analysis techniques are used.
2. Grounded theory (GT) analysis: in which there is no mention of the grounded theory approach or methodology. Grounded theory is only associated with the data analysis techniques.
3. Other: in which there is no mention of grounded theory and the analysis approach has been called ‘coding’.

We examined the data gathering techniques used in the 59 papers. This enabled us to identify the type of data being analysed using OAS coding. Table 2-13 illustrates the list of possible data gathering techniques and the percentages of the 59 papers that used them. In some cases, the papers reported using more than one data gathering technique.

Table 2-13 Analysis of the data gathering techniques used in the 59 papers.

Data gathering technique	Usage percentage	Examples
Interviews	87%	Pauleen, 2003; Goulielmos, 2004; Kirsch & Haney, 2006; Smolander et al., 2008; Chakraborty et al., 2010; O'Reilly & Finnegan, 2010; Gleasure, 2015
Documentation	44%	Ryan & Valverde, 2006; Keil et al., 2007; Goode & Gregor, 2009; Maldonado, 2010
Observations	23%	Huang et al., 2001; Vaast & Walsham, 2013

The majority of the 59 papers rely on interviews as the primary data gathering technique, while documentation is the second most commonly used technique. Many of the papers analysed apply both interviews and observations and some include case documents. However, some papers rely solely on analysing documents and are considered to be literature analyses (such as Goode and Gregor, 2009).

In terms of types of analysis, the 59 papers are distributed across the three types of data analysis context (see Figure 2-7). Most of the papers come under the GT approach. However, more than one-third refer to the coding techniques as GT analysis and apply them in different contexts. This is an indication of the use of the three coding techniques in a context of non-grounded theory research. In addition, there are 13 papers in which there is no mention of grounded theory, either in their approach or analysis, but which used the three coding techniques referred to by Strauss and Corbin (such as Maldonado, 2010 and Kane and Labianca, 2011).

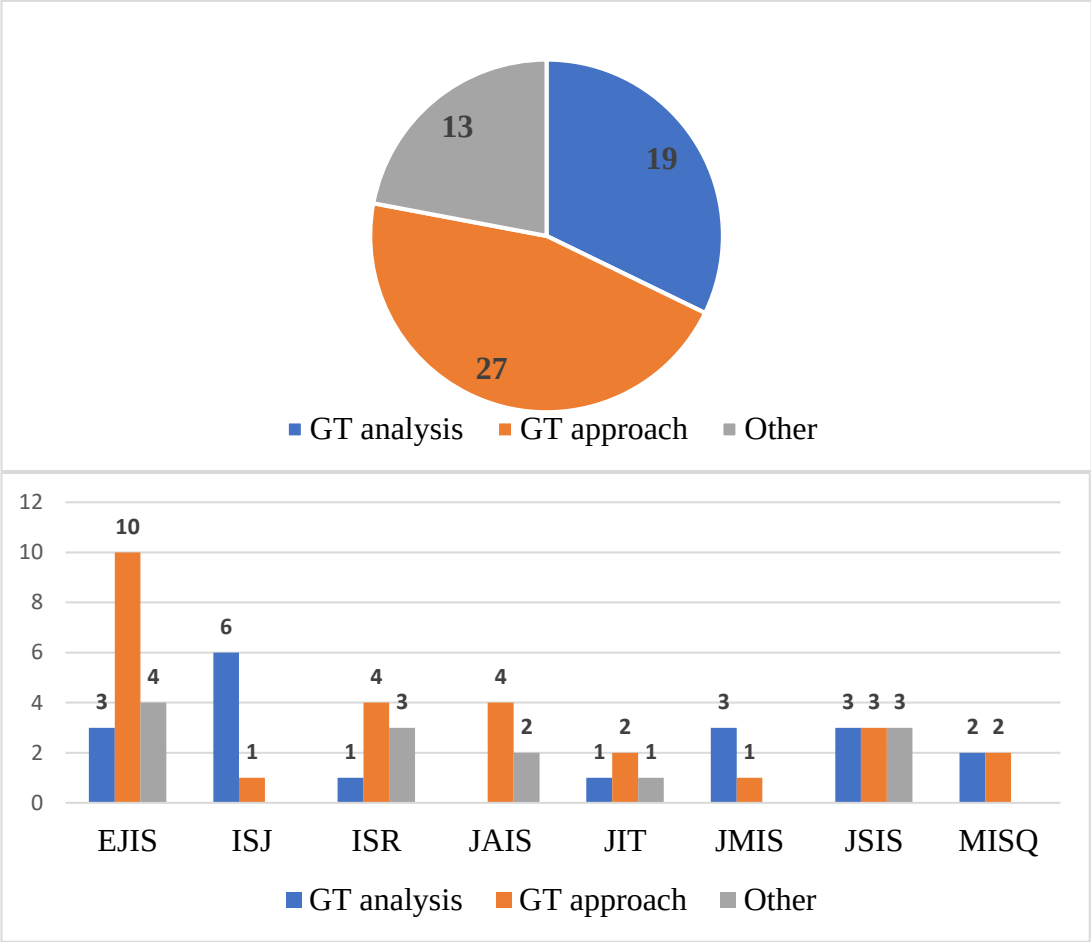

Figure 2-7 Context of the coding techniques.

Our findings show that the majority of the papers that used OAS coding did so in the context of a grounded theory approach. These 27 papers explicitly follow the grounded theory approach which incorporate the three coding techniques and are aimed at building a theory, such as Galal (2001), Huang et al., (2001) and Day et al. (2009).

In addition, 19 papers did not follow a grounded theory approach or research methodology, although they used OAS coding and referred to them as grounded theory analysis. These papers mainly follow theory building approach by referring

to another method, such as case study research (c.f. Yin, 2003), building theory from case studies (c.f. Eisenhardt, 1985) or action research (c.f. Susman & Evered, 1978). Hence, OAS coding can be used in different contexts for research approaches that are aimed at theory building.

The final classification of the papers is 'Other'. The 13 papers in this classification used the OAS coding techniques. However, these papers do not mention grounded theory, either as an approach or a data analysis technique. It can be argued that some of these papers are aimed at empirically building a theory (such as Tan et al., 2015). However, there are some papers that have used the coding techniques in order to test a theory empirically, such as Maldonado (2010) and Chan et al. (2011). This indicates the power of using the coding techniques for non-theory-building research.

The following section presents a deeper analysis of how these papers operationalised open, axial and selective coding techniques.

2.2.6.3 Operationalising OAS coding

Taking into consideration the 59 papers, we analysed the research methodology section of each of the papers in depth. We did this in order to code all the concepts related to the three coding techniques to enable us to clarify how the coding was conducted. Initially, the 59 papers were coded according to the coding framework, which includes the three coding techniques and the activities involved (see Figure 2-5). We were able to classify the papers on a scale with (Explicit) material at one end and reference-only material (Hints) at the other end (c.f. Seidel & Urquhart,

2013). ‘Explicit’ material indicates papers that explicitly mention the three coding techniques as stated in the coding framework and fully explain the operationalising processes, whereas ‘Hints’ indicates papers that only refer to the three coding techniques without specifying the processes or the activities involved.

Table 2-14 shows the classification of the 59 papers according to the coding framework. There are three possible scenarios for each of the coding activities:

1. Activities explicitly explained. In this case, a tick is placed under the activity.
2. Mentioning the activity but with a different meaning, such as ‘discovering categories’, which is in many cases considered as an activity for axial coding, instead of ‘Linking categories in sets of relationships’. In this case, therefore, an activity is left blank.
3. Following Strauss and Corbin without specifying (hints); the columns for these papers are left blank.

Table 2-14 Classifications of the 59 papers by coding framework.

	Classification Number	Reference	Number of papers	Open coding		Axial coding		Selective coding	
				Labelling concepts	Discovering categories	Linking categories in sets of relationships	Use paradigm model	Selection of core categories	Identifying the story
Explicit ⟶ Hints	1	Galal, 2001; Huang et al., 2001; Work, 2002; Baskerville & Pries-Heje, 2004; Day et al., 2009; Palka et al., 2009; Maldonado, 2010	7	√	√	√	√	√	√
	2	Keil et al., 2007; Xu & Ramesh, 2007; Goode & Gregor, 2009; Ramesh et al., 2010; Walsh et al., 2010; Chang et al., 2011; Kane & Labianca, 2011; Clemmensen, 2012; Mattarelli et al., 2013; Morgan et al., 2013	10	√	√	√		√	√
	3	Strong & Volkoff, 2010; Zahedi & Bansal, 2011; Williams & Karahanna, 2013	3	√		√	√	√	√
	4	Feller et al., 2008; O'Reilly & Finnegan, 2010	2	√	√	√	√		
	5	Kock, 2001; Goulielmos, 2004; Berente et al., 2011; Berente & Yoo, 2012; Vaast & Walsham, 2013; Huang et al., 2014; Strong et al., 2014; Gleasure, 2015; Karoui et al., 2015	9	√		√		√	√
	6	Kirsch, 2004; Smolander et al., 2008; Petrini & Pozzebon, 2009; Feller et al., 2012; Bagayogo et al., 2014	5	√	√	√			
	7	Kirsch, 1997; Pauleen & Yoong, 2001; Kirsch & Haney, 2006; Berente et al., 2010; Chan et al., 2011; Holmström & Sawyer, 2011; Tallon et al., 2013; Tan et al., 2015	8	√				√	√
	8	Pauleen, 2003; Abraham et al., 2013	2	√		√	√		
	9	Ransbotham & Mitra, 2009; Lederman & Johnston, 2011; Leonardi, 2014	3	√		√			
	10	Lindgren et al., 2004; Ryan & Valverde, 2006; Hackney et al., 2007; Vaast, 2007; Butler & Murphy, 2008; Matsuo et al., 2008; Chakraborty et al., 2010; Jenkin & Chan, 2010; Vannoy & Salam, 2010; Butler, 2011	10						

From Table 2-14, it can be seen that 24 papers (see classifications 1,2,4, and 6) conduct open coding for the purpose of ‘labelling concepts’ and ‘discovering categories’ from data in order for these to be used for axial coding, such as O'Reilly and Finnegan (2010) and Morgan et al. (2013). This can be considered the ideal method of using open coding. However, different terms are used to refer to concepts and categories. For example, one paper (Smolander et al., 2008) calls the concepts ‘seed categories’ and the categories ‘super categories’, which reflects the original meaning of the concepts and categories in open coding. In open coding, the concepts and categories emerge from the data. In terms of labelling concepts and categories, it is also recommended that these be taken from the actual named in the data where possible (c.f. Petrini & Pozzebon, 2009; Strong & Volkoff, 2010), although predefined categories from the literature can also be used, as in Huang et al. (2014).

From Table 2-14, we can also observe that 22 papers (see classifications 5,7,8, and 9) only use open coding to label concepts. The majority of these papers overlap between open and axial coding and discovering categories was considered in the axial coding phase. For example, Kirsch (1997), Holmström and Sawyer (2011) and Abraham et al. (2013) label concepts during open coding, then compare the concepts for similarities and differences in order to discover higher-level categories during the axial coding stage. The original open coding by presented Strauss and Corbin considers this action to be part of the open coding stage. Therefore, we can see how the operationalisation causes confusion around coding.

In terms of axial coding, our analysis shows 14 papers (see classifications 1,3,4, and 8) are considered to demonstrate the ideal operationalisation of axial coding, as they have linked categories in sets of relationships as well as using a paradigm model as their main focus in axial coding. However, we observed that, in the axial coding stage, the majority of the papers consider that the linking of categories in sets of relationships is the main action that should be conducted, without specifying the paradigm model, such as Goulielmos (2004), Berente and Yoo (2012) and Strong et al. (2014). These papers have lost the advantages and the flavour of axial coding.

In terms of applying paradigm models, the majority of the papers do not mention a paradigm model at all, even if they have followed Strauss and Corbin for axial coding. We found that only 15 papers mention a paradigm model and apply the concepts of forming relationships between categories. The remainder of the papers do not mention a paradigm model, which shows a lack of understanding in operationalising axial coding as it was not fully conducted. Axial coding should be associated with a paradigm model, which is the main distinction between Strauss and Corbin (1990) and Glaser (1992) in terms of coding techniques, this argument also is made by (Seidel & Urquhart, 2013).

In addition, our analysis shows that not all of a paradigm model's elements have to be used during axial coding. For example, Chakraborty et al. (2010) and Williams and Karahanna (2013) use elements of their 'own paradigm model', whereas, in contrast, Day et al. (2009) and Strong and Volkoff (2010) use the same elements of the original paradigm model as prescribed by Strauss and Corbin (1990).

In terms of their use of selective coding, the majority of the papers (37) (see classifications 1,2,3,5, and 7) undertake the selection of core categories and identify a story as stated in the coding framework, such as Keil et al. (2007), Day et al. (2009) and Morgan et al. (2013), which can be considered as ideal use of selective coding. However, some of the papers, such as Feller et al. (2008) and O'Reilly and Finnegan (2010), conduct selective coding simply by re-doing axial coding and identifying causal relationships between categories. Others, such as Huang et al. (2001) and Kirsch (2004), re-do open coding by generating categories or comparing them in terms of similarities and differences.

Finally, there are a number of papers that do not specify any of the three coding techniques or are missing a part (see classification 10), such as Lindgren et al. (2004), Ryan and Valverde (2006) and Jenkin and Chan (2010). These papers explicitly mention that they 'are following Strauss and Corbin's coding techniques', although there is no further explanation of how they used the coding. This is, therefore, considered an implicit (hints) use of coding techniques.

2.2.7 Discussion and recommendations for novice researchers

In this section, we discuss the findings from our analysis of the 59 coded papers in order to clarify some of the confusion regarding the use of OAS coding that exists. We argue that this confusion is not helpful for novice researchers. Initially, we found that the use of OAS coding is increasing within IS academic publications. However, we observe different usage patterns and flexibility in using open, axial and selective coding. From our point of view, these differences are driven by the

nature of the research strategy; however, it can also be considered a misunderstanding of how to operationalise the OAS coding techniques themselves.

In addition, we observe that many of the coded papers have not explained how they used the coding techniques, which creates more confusion for novice researchers. However, some of the papers are more explicit and comprehensive in explaining the process of operationalising the coding of their data. We argue that these papers constitute a list of recommended readings for novice researchers. Therefore, our recommendation for novice researchers is to review further the seven papers (top row of Table 2-14) that explicitly explain the use of the three coding techniques. Thereafter, novice researchers can review those papers that have explained one or more of the activities for the three coding techniques.

In terms of open coding, we found that the primary tasks to be undertaken are labelling the concepts based on the data, then discovering categories by comparing the concepts and creating an abstracted layer of similar concepts. Labelling the concepts should be driven by the research objectives and what the data are revealing, and hence is considered to be the seed for the resulting data and helping to make sense of the data collected. However, the process of labelling concepts and discovering categories is iterative, as, in most cases, the concepts emerge from the data, which helps to shape the final categories. Therefore, we argue that this type of coding requires a high level of coding skill.

In axial coding, the categories that result from open coding are compared for similarities and differences in order to make causal relationships between them. Here, we found many of the researchers misunderstand this kind of activity by only

re-grouping categories at a higher level, whereas they should be related to causal relationships. In many cases, it requires the researcher to return to the original text (the data collected) of the categories and the associated concepts in order to understand the actual relationships. Therefore, we argue that having a paradigm model helps to operationalise the right approach to axial coding (see Figure 2-8, an example of the relationships between the elements of the paradigm model). In fact, not being able to visualise what constitutes the workings of a paradigm model might also explain why it is not that prominent a feature of axial coding in the papers reviewed.

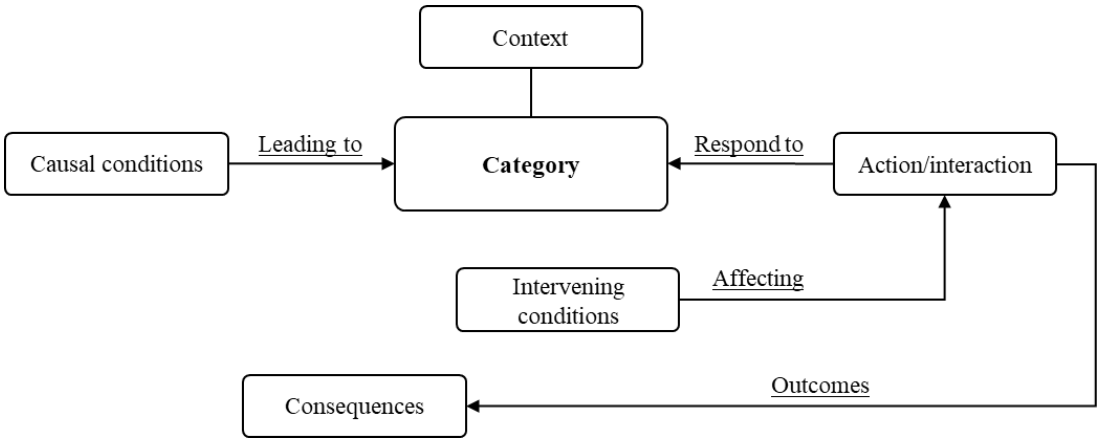

Figure 2-8 Example of the relationships between the paradigm model's elements.

In terms of a paradigm model, we found that, although there are suggested elements for such a model to help shape the causal relationships between categories, researchers can create a paradigm to fit their own research objectives. The relationships between the categories within the paradigm model should be shaped from the actual data, and can also be shaped from the literature in some cases.

Selective coding aims to form a theoretical framework by selecting core categories (the central phenomenon of the theoretical model) which result from the paradigm

model. Then, as a final stage, after comparing the core categories with the original data, a story is built that fulfils the research objectives and answers the research questions, if any. This is achieved by stating with what the data reveal after the coding. Therefore, we argue that selective coding is used as a final stage to form the final theoretical framework.

Finally, Figure 2-9 shows the OAS coding framework with our recommendations for novice researchers under each of the activities.

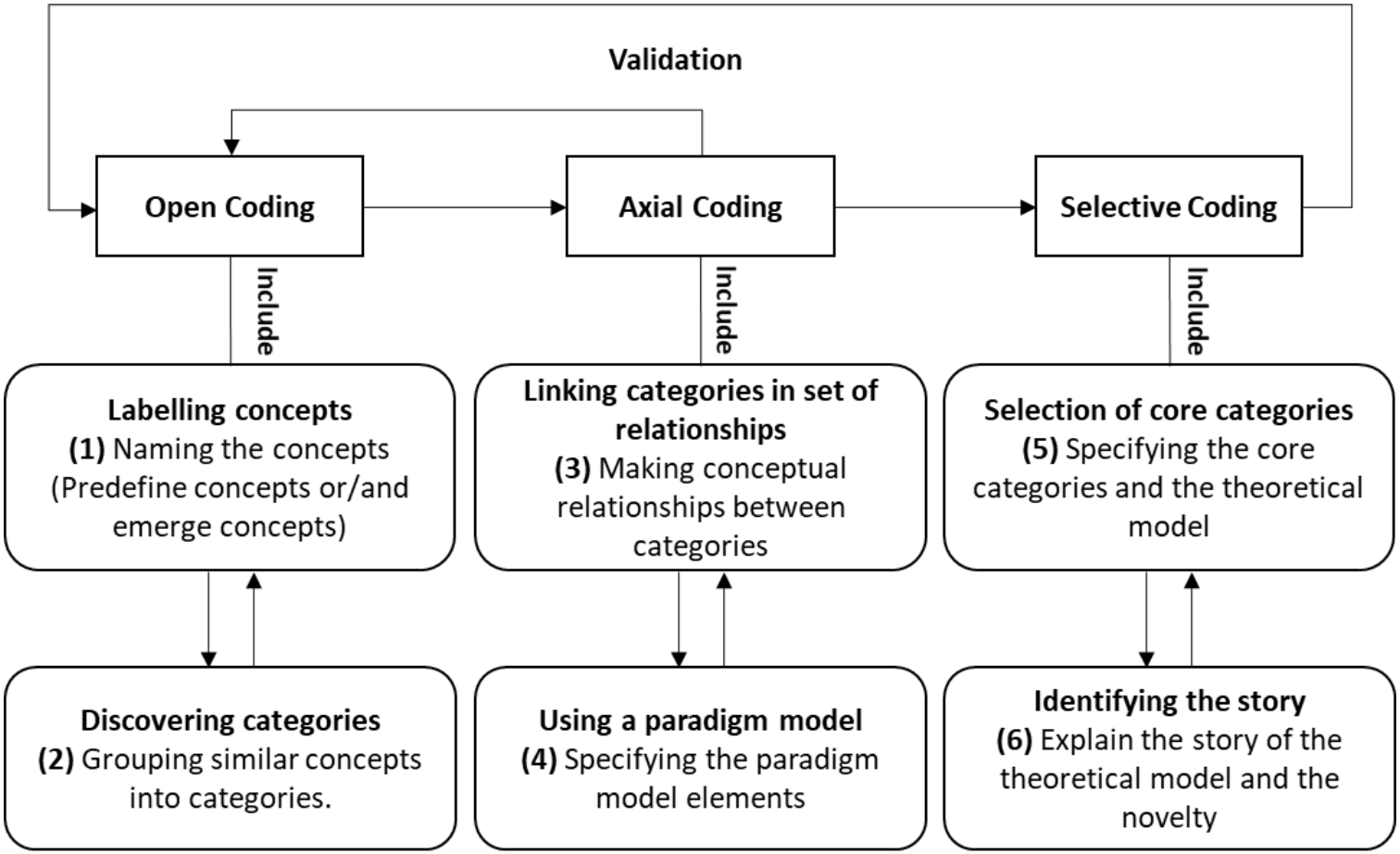

Figure 2-9 The OAS Coding Framework.

From the OAS coding framework in Figure 2-9, it can be seen that the three coding techniques are executed iteratively rather than in a linear fashion. The activities included within each coding technique are also iterative. For example, researchers can start discovering categories after labelling some of the concepts, although the activity of labelling concepts is still being developed. This iterative manner is applicable across all the activities within OAS coding.

In open coding, it is recommended to decide to use predefined concepts which are driven by either the literature or a research model, or labelling the concepts incrementally from the data. In some cases, the researcher can have a list of predefined concepts and have new concepts emerge from the data. As a second activity within open coding, these concepts should be compared and those that are similar grouped into categories. Naming the categories can involve using either one of the concept labels or a name that reflects the meaning of the concepts contained within the category.

In axial coding, the categories that are discovered can be linked in the form of relationships using the concept of a paradigm model in which causal relationships are the initial step in linking the categories. Here, the researcher is required to review the concepts contained in each category in order to form better relationships. Although the paradigm model has certain elements that can be used to classify the categories, the researcher can select the most useful element of the paradigm model, or make his or her own paradigm model, that has elements driven by the research model.

Finally, after comparing the categories and identifying the relationships during the selective coding stage, the researcher should select core categories that fulfil the research objective or answer the research questions, if any. By selecting core categories, the researcher will be able to build a theoretical model and then explain the model according to the categories it contains in order to provide a narrative and highlight the novelty of the research.

2.2.8 Concluding remarks

According to Rowe (2014) there is a need within the IS community to publish more literature reviews. He argues that "*literature reviews can be highly valuable*" and "*every researcher looks for [a literature review] when starting a research study*" (Rowe, 2014, p.242). So where the main goal of a literature review is "*to classify what has been produced by the literature*" (Rowe, 2014, p.243) we believe that we have achieved this for OAS coding techniques and mapped the territory (see Table 2-14) using our OAS coding Framework (see Figure 2-9).

Rowe (2014, p.246) suggests that "*the quality of a literature review depends on its systematicity, since systematicity implies reproducibility through documenting the search process and potentially indicates comprehensiveness*". This research study identified and analysed 59 published IS studies that have operationalised OAS coding techniques. Using a systematic approach, through the eight coding steps of content analysis, the selection process yielded 192 publications that were subjected to selection and exclusion criteria, which led to the exclusion of 25. Following a more in-depth review of the remaining 167 publications, 59 were found to serve the research purpose explicitly. These 59 publications were analysed using an open

coding analysis technique suggested by Corbin and Strauss (1990). This technique was selected to conduct an in-depth content analysis of the OAS coding techniques mentioned in these publications. Therefore, we believe that we have achieved the systematicity required to ensure the reproducibility of our work by others.

The motivation for writing this paper is that the three coding techniques to which it refers are a powerful means of analysing qualitative data. However, our analysis clearly shows that the descriptions of the three coding techniques used by IS scholars have been rather vague. We observed many overlapping activities between the three coding techniques. For example, some papers conducted axial coding with open coding activities. The majority of the papers also missed the core value of the paradigm model during axial coding. Our analysis also shows that there is uncertainty in conducting the three coding techniques. This causes confusion for novice researchers when they come to conduct the three coding techniques.

We have, therefore, listed papers we recommend (Galal, 2001; Huang et al., 2001; Work, 2002; Baskerville & Pries-Heje, 2004; Day et al., 2009; Palka et al., 2009; Maldonado, 2010), as they are explicit and comprehensive in the way they conduct the three coding techniques in different research contexts. This list of papers is intended to act as a recommended reading list (the must reads) for novice researchers or indeed any researcher looking to execute a qualitative data analysis strategy. In fact, we can go one step further and recommend that all research methods modules, associated with graduate research programmes, incorporate these seven papers as part of their reading lists. Indeed, a further invaluable teaching and learning exercise would be for novice researchers to compare the papers classified

in row 1 of Table 2-14 with those classified in row 10; thereby creating a compare and contrast of the ‘explicit’ and the ‘hints’ approaches used to narrate the operationalisation of OAS coding techniques.

We found that a coding framework can be useful in following appropriate activities when operationalising coding. The list of recommended papers was analysed to enable us to identify specifications for each of the activities in order to create a roadmap that would enable novice researchers to undertake the three coding techniques (see Figure 2-9). The list of the remaining papers in Table 2-14 can also be used for further understanding aspects of the use of coding techniques. Finally, our analysis indicates that the coding techniques can be useful for a variety of research activities and can readily be adapted to answer specific questions.

2.2.9 References

ABRAHAM C, BOUDREAU M, JUNGLAS I and WATSON R (2013) Enriching our theoretical repertoire: The role of evolutionary psychology in technology acceptance. European Journal of Information Systems, 22(1), 56-75.

BAGAYOGO FF, LAPOINTE L and BASSELLIER G (2014) Enhanced use of IT: A new perspective on post-adoption. Journal of the Association for Information Systems, 15(7), 361-387.

BASKERVILLE R and PRIES-HEJE J (2004) Short cycle time systems development. Information Systems Journal, 14(3), 237-264.

BENBUNAN-FICH, R. and BENBUNAN, A., 2007. Understanding user behavior with new mobile applications. The Journal of Strategic Information Systems, 16(4), pp. 393-412.

BERENTE N, GAL U and YOO Y (2010) Dressage, control, and enterprise systems: The case of NASA's full cost initiative. European Journal of Information Systems, 19(1), 21-34.

BERENTE N, HANSEN S, PIKE JC and BATEMAN PJ (2011) Arguing the value of virtual worlds: Patterns of discursive sensemaking of an innovative technology. MIS Quarterly, 35(3), 685-709.

BERENTE N and YOO Y (2012) Institutional contradictions and loose coupling: Post Implementation of NASA's enterprise information system. Information Systems Research, 23(2), 376-396.

BERG, B. L. (2001). Methods for the social sciences Pearson Education Inc, United States of America.

BHATTACHERJEE, A., 2012. Social science research: principles, methods, and practices. Textbooks Collection. Book 3.

BIRKS DF, FERNANDEZ W, LEVINA N and NASIRIN S (2013) Grounded theory method in information systems research: Its nature, diversity and opportunities. European Journal of Information Systems, 22(1), 1-8.

BRYMAN, A., 1984. The Debate about Quantitative and Qualitative Research: A Question of Method or Epistemology? The British journal of sociology, 35(1), pp. 75-92.

BUTLER T (2011) Compliance with institutional imperatives on environmental sustainability: Building theory on the role of green IS. The Journal of Strategic Information Systems, 20(1), 6-26.

BUTLER T and MURPHY C (2008) An exploratory study on IS capabilities and assets in a small-to-medium software enterprise. Journal of Information Technology, 23(4), 330-344.

CHAKRABORTY S, SARKER S and SARKER, S (2010) An exploration into the process of requirements elicitation: A grounded approach. Journal of the Association for Information Systems, 11(4), 212-249.

CHAN CML, HACKNEY R, PAN SL and CHOU T (2011) Managing e-government system implementation: A resource enactment perspective. European Journal of Information Systems, 20(5), 529-541.

CHANG CLH, CHEN V, KLEIN G and JIANG JJ (2011) Information system personnel career anchor changes leading to career changes. European Journal of Information Systems, 20(1), 103-117.

CLEMMENSEN T (2012) Usability problem identification in culturally diverse settings. Information Systems Journal, 22(2), 151-175.

DAY JM, JUNGLAS I and SILVA L (2009) Information flow impediments in disaster relief supply chains. Journal of the Association for Information Systems, 10(8), 637-660.

DEZDAR, S. and SULAIMAN, A., 2009. Successful enterprise resource planning implementation: taxonomy of critical factors. Industrial Management & Data Systems, 109(8), pp. 1037-1052.

EAVES, Y. D. (2001). A synthesis technique for grounded theory data analysis. Journal of Advanced Nursing, 35(5), 654-663.

EISENHARDT, K. M. (1989). Building theories from case study research. Academy of management review, 14(4), 532-550.

FELDMAN SS and HORAN TA (2011) The dynamics of information collaboration: A case study of blended it value propositions for health information exchange in disability determination. Journal of the Association for Information Systems, 12(2), 189-207.

FELLER, J, FINNEGAN P, HAYES J and O'REILLY P (2012) 'Orchestrating' sustainable crowdsourcing: A characterisation of solver brokerages. The Journal of Strategic Information Systems, 21(3), 216-232.

FELLER J, FINNEGAN P, FITZGERALD B and HAYES J (2008) From peer production to productization: A study of socially enabled business exchanges in open source service networks. Information Systems Research, 19(4), 475-493.

FINNEY, S. and CORBETT, M., 2007. ERP implementation: a compilation and analysis of critical success factors. Business Process Management Journal, 13(3), pp. 329-347.

FURNEAUX B and WADE MR (2011) An exploration of organizational level information systems discontinuance intentions. MIS Quarterly, 35(3) , 573-598.

GALAL G (2001) From contexts to constructs: The use of grounded theory in operationalising contingent process models. European Journal of Information Systems, 10(1), 2-14.

GIVEN, L. M. (2008). The sage encyclopedia of qualitative research methods Sage Publications.

GLASER, B.G., 1992. Emergence vs forcing: Basics of grounded theory analysis. Sociology Press.

GLEASURE R (2015) Resistance to crowdfunding among entrepreneurs: An impression management perspective. The Journal of Strategic Information Systems, 24(4), 219-233.

GOO, J., KISHORE, R., RAO, H.R. and NAM, K., 2009. The role of service level agreements in relational management of information technology outsourcing: an empirical study. MIS Quarterly, , pp. 119-145.

GOODE S and GREGOR S (2009) Rethinking organisational size in IS research: Meaning, measurement and redevelopment. European Journal of Information Systems, 18(1), 4-25.

GOULIELMOS M (2004) Systems development approach: Transcending methodology. Information Systems Journal, 14(4), 363-386.

GRAHLMANN, K.R., HELMS, R.W., HILHORST, C., BRINKKEMPER, S. and VAN AMERONGEN, S., 2012. Reviewing enterprise content management: A functional framework. European Journal of Information Systems, 21(3), pp. 268-286.

HACKNEY R, JONES S and LÖSCH A (2007) Towards an e-government efficiency agenda: The impact of information and communication behaviour on e-reverse auctions in public sector procurement. European Journal of Information Systems, 16(2), 178-191.

HEATH, H., & COWLEY, S. (2004). Developing a grounded theory approach: A comparison of glaser and strauss. International Journal of Nursing Studies, 41(2), 141-150.

HOLMSTRÖM J and SAWYER S (2011) Requirements engineering blinders: Exploring information systems developers' black-boxing of the emergent character of requirements. European Journal of Information Systems, 20(1), 34-47.

HUANG J, MAKOJU E, NEWELL S and GALLIERS RD (2003) Opportunities to learn from 'failure' with electronic commerce: a case study of electronic banking. Journal of Information Technology, 18(1), 17-26.

HUANG JC, NEWELL S and PAN SL (2001) The process of global knowledge integration: A case study of a multinational investment bank's Y2K program. European Journal of Information Systems, 10(3), 161-174.

HUANG P, PAN SL and OUYANG TH (2014) Developing information processing capability for operational agility: Implications from a chinese manufacturer. European Journal of Information Systems, 23(4), 462-480.

JENKIN TA and CHAN YE (2010) IS project alignment–a process perspective. Journal of Information Technology, 25(1), 35-55.

JOHNSON, R.B. and ONWUEGBUZIE, A.J., 2004. Mixed methods research: A research paradigm whose time has come. Educational researcher, 33(7), pp. 14-26.

KANE GC and LABIANCA G (2011) IS avoidance in health-care groups: A multilevel investigation. Information Systems Research, 22(3), 504-522.

KAPLAN, B. and DUCHON, D., 1988. Combining qualitative and quantitative methods in information systems research: a case study. MIS quarterly, , pp. 571-586.

KAROUI M, DUDEZERT A and LEIDNER DE (2015) Strategies and symbolism in the adoption of organizational social networking systems. The Journal of Strategic Information Systems, 24(1), 15-32.

KEIL M, IM GP, and MÄHRING M (2007) Reporting bad news on software projects: The effects of culturally constituted views of face-saving. Information Systems Journal, 17(1), 59-87.

KIRSCH LJ (2004) Deploying common systems globally: The dynamics of control. Information Systems Research, 15(4), 374-395.

KIRSCH LJ and HANEY MH (2006) Requirements determination for common systems: Turning a global vision into a local reality. The Journal of Strategic Information Systems, 15(2), 79-104.

KIRSCH LS (1997) Portfolios of control modes and IS project management. Information Systems Research, 8(3), 215-239.

KOCK N (2001) Asynchronous and distributed process improvement: The role of collaborative technologies. Information Systems Journal, 11(2), 87-110.

LEDERMAN R and JOHNSTON RB (2011) Decision support or support for situated choice: Lessons for system design from effective manual systems. European Journal of Information Systems, 20(5), 510-528.

LEONARDI PM (2014) Social media, knowledge sharing, and innovation: Toward a theory of communication visibility. Information Systems Research, 25(4), 796-816.

LEONARDI, P.M., 2013. When does technology use enable network change in organizations? A comparative study of feature use and shared affordances. MIS Quarterly, 37(3) , 749-775.

LEVINA, N. and ROSS, J.W., 2003. From the vendor's perspective: exploring the value proposition in information technology outsourcing. MIS Quarterly, , pp. 331-364.

LINDGREN R, HENFRIDSSON O and SCHULTZE U (2004) Design principles for competence management systems: A synthesis of an action research study. MIS Quarterly, 28(3), 435-472.

MALDONADO E (2010) The process of introducing FLOSS in the public administration: The case of Venezuela. Journal of the Association for Information Systems, 11(11), 756-783.

MATAVIRE R and BROWN I (2013) Profiling grounded theory approaches in information systems research. European Journal of Information Systems, 22(1), 119-129.

MATSUO M, WONG CWY and LAI K (2008) Experience-based learning of Japanese IT professionals: A qualitative research. The Journal of Strategic Information Systems, 17(3), 202-213.

MATTARELLI E, BERTOLOTTI F and MACRÌ DM (2013) The use of ethnography and grounded theory in the development of a management information system. European Journal of Information Systems, 22(1), 26-44.

MORGAN L, FELLER J and FINNEGAN P (2013) Exploring value networks: Theorising the creation and capture of value with open source software. European Journal of Information Systems, 22(5), 569-588.

MIRANDA, S. M., KIM, I., & SUMMERS, J. D. (2015). Jamming with social media: How cognitive structuring of organizing vision facets affects IT innovation diffusion. MIS Quarterly, 39(3)

MYERS, M.D., 1997. Qualitative research in information systems. Management Information Systems Quarterly, 21, pp. 241-242.

MYERS, M. (2009) Qualitative Research in Business and Management, Sage, UK.

O'REILLY P and FINNEGAN P (2010) Intermediaries in inter-organisational networks: Building a theory of electronic marketplace performance. European Journal of Information Systems, 19(4), 462-480.

OLSSON HH, CONCHÚIR EÓ, ÅGERFALK PJ and FITZGERALD B (2008) Two-stage offshoring: An investigation of the irish bridge. MIS Quarterly, 32(2) , 257-279.

PALKA W, POUSTTCHI K and WIEDEMANN DG (2009) Mobile word-of-mouth–A grounded theory of mobile viral marketing. Journal of Information Technology, 24(2), 172-185.

PAULEEN DJ (2003) An inductively derived model of leader-initiated relationship building with virtual team members. Journal of Management Information Systems, 20(3), 227-256.

PAULEEN DJ and YOONG P (2001) Relationship building and the use of ICT in boundary-crossing virtual teams: A facilitator's perspective. Journal of Information Technology, 16(4), 205-220.

PETRINI M and POZZEBON M (2009) Managing sustainability with the support of business intelligence: Integrating socio-environmental indicators and organisational context. The Journal of Strategic Information Systems, 18(4), 178-191.

RAMESH B, CAO L and BASKERVILLE R (2010) Agile requirements engineering practices and challenges: An empirical study. Information Systems Journal, 20(5), 449-480.

RANSBOTHAM S and MITRA S (2009) Choice and chance: A conceptual model of paths to information security compromise. Information Systems Research, 20(1), 121-139.

ROWE, F. (2014) What literature review is not: diversity, boundaries and recommendations, European Journal of Information Systems, 23(3), pp.241-255.

RYAN G and VALVERDE M (2006) Waiting in line for online services: A qualitative study of the user's perspective. Information Systems Journal, 16(2), 181-211.

SCOTT JE (2000) Facilitating interorganizational learning with information technology. Journal of Management Information Systems, 17(2), 81-113.

SEIDEL S and URQUHART C (2013) On emergence and forcing in information systems grounded theory studies: The case of strauss and corbin. Journal of Information Technology, 28(3), 237-260.

SHANG, S. and SEDDON, P.B., 2002. Assessing and managing the benefits of enterprise systems: the business manager's perspective. Information systems journal, 12(4), pp. 271-299.

SMOLANDER K, ROSSI M and PURAO S (2008) Software architectures: Blueprint, literature, language or decision? European Journal of Information Systems, 17(6), 575-588.

STRAUSS, A.L. AND CORBIN, J. (1990). Basics of Qualitative Research, 1st edn, Thousand Oaks, CA: Sage.

STRAUSS, A.L. AND CORBIN, J. (1998). Basics of Qualitative Research: Techniques and procedures for developing grounded theory, 2nd edn, London: Sage

STRAUSS, A.L. AND CORBIN, J. (2008). Basics of Qualitative Research, Thousand Oaks, CA: Sage

STRONG DM, JOHNSON SA, BENGISU T, TRUDEL J, and VOLKOFF O (2014) A theory of organization-EHR affordance actualization. Journal of the Association for Information Systems, 15(2), 53-85.

STRONG DM and VOLKOFF O (2010) Understanding Organization—Enterprise system fit: A path to theorizing the information technology artifact. MIS Quarterly, 34(4) , 731-756.

SUSMAN, G. I., and EVERED, R. D. (1978). An assessment of the scientific merits of action research. Administrative science quarterly, 582-603.

SUTTON, R.I. and STAW, B.M., 1995. What Theory is Not. Administrative Science Quarterly, 40(3), pp. 371-384.

TALLON PP, RAMIREZ RV and SHORT JE (2013) The information artifact in IT governance: Toward a theory of information governance. Journal of Management Information Systems, 30(3), 141-178.

TAN B, PAN SL, LU X and HUANG L (2015) The role of IS capabilities in the development of multi-sided platforms: The digital ecosystem strategy of Alibaba. com. Journal of the Association for Information Systems, 16(4), 248-280.

URQUHART C, LEHMANN H and MYERS MD (2010) Putting the 'theory' back into grounded theory: Guidelines for grounded theory studies in information systems. Information Systems Journal, 20(4), 357-381.

URQUHART C and FERNANDEZ W (2013) Using grounded theory method in information systems: The researcher as blank slate and other myths. Journal of Information Technology, 28(3), 224-236.

VAAST E (2007) Danger is in the eye of the beholders: Social representations of information systems security in healthcare. The Journal of Strategic Information Systems, 16(2), 130-152.

VAAST E and WALSHAM G (2013) Grounded theorizing for electronically mediated social contexts. European Journal of Information Systems, 22(1), 9-25.

VANNOY SA and SALAM A (2010). Managerial interpretations of the role of information systems in competitive actions and firm performance: A grounded theory investigation. Information Systems Research, 21(3), 496-515.

WALSH I, KEFI H and BASKERVILLE R (2010) Managing culture creep: Toward a strategic model of user IT culture. The Journal of Strategic Information Systems, 19(4), 257-280.

WALSHAM, G., 2006. Doing interpretive research. European journal of information systems, 15(3), pp. 320-330.

WEBB B and MALLON B (2007) A method to bridge the gap between breadth and depth in IS narrative analysis. Journal of the Association for Information Systems, 8(7), 368-381.

WEBSTER J and WATSON RT (2002) Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly, 26(2), xiii-xxiii.

WILLIAMS CK and KARAHANNA E (2013) Causal explanation in the coordinating process: A critical realist case study of federated IT governance structures. MIS Quarterly, 37(3), 933-964.

WOLFSWINKEL JF, FURTMUELLER E and WILDEROM CP (2013) Using grounded theory as a method for rigorously reviewing literature. European Journal of Information Systems, 22(1), 45-55.

WORK B (2002) Patterns of software quality management in TickIT certified firms. European Journal of Information Systems, 11(1), 61-73.

XU P and RAMESH B (2007) Software process tailoring: An empirical investigation. Journal of Management Information Systems, 24(2), 293-328.

YIN, R. K. (2003). Case study research: Design and methods. Sage publications

ZAHEDI FM and BANSAL G (2011) Cultural signifiers of web site images. Journal of Management Information Systems, 28(1), 147-200.

2.3 Paper 3

Critical Success Factors for Data Governance: A Theory Building Approach

2.3.1 Abstract

The objective of this research study is to present the critical success factors (CSFs) for data governance (DG). This paper reports on a single case study where data are gathered through semi-structured interviews following the CSF approach and analysed by applying open, axial, and selective coding techniques. The findings of this research are presented as seven CSFs, which are ranked in order of importance. Based on our case analysis, employee data competencies were the most important factor for data governance. Furthermore, we highlight the relationships between the CSFs in order to understand their possible interconnectedness.

2.3.2 Keywords:

Data governance, CSFs, case study, open coding, axial coding, selective coding

2.3.3 Introduction

Thinking about data strategically is a problem for many organisations today. Governing data has become vital in running a business successfully, in order for

data to be treated as a valuable asset (Khatri & Brown, 2010; Otto, 2015). In recent years, the volume of data used within organisations has increased dramatically, playing a critical role in business operations (Tallon, Ramirez, & Short, 2013). In particular, data influence both operational and strategic decisions. It is argued that a lack of trust in data can lead to the wasting of up to 50% of knowledge workers' time, spent "*hunting for data*" (Redman, 2013, p. 4), whereas, when "*data is trusted, it gets shared*", which can drive higher returns on data investments (Information Builders, 2014, p. 8). Hence, the question arises: how do we ensure we are building trusted data? A recent study by Holt, Ramage, Kear, and Heap (2015) indicated that 45% of the participants, who were from the global community of data-base and data professionals, did not have data governance policies in place. Hence, data governance requires more attention from stakeholders.

Although data governance is considered to be a relatively emerging subject (Kamioka, Luo, & Tapanainen, 2016; Rasouli, Eshuis, Trienekens, Kusters, & Grefen, 2016), several researchers have proposed different data governance models (c.f. Khatri & Brown, 2010; Otto, 2011; Panian, 2010; Weber et al., 2009; Wende, 2007). These researchers help in our understanding of the data governance subject and in shaping its boundaries. However, only a limited number of papers have examined the critical success factors (CSFs) for data governance. This research aims to contribute to the body of knowledge by inductively identifying the CSFs for data governance, following the building theory from case study research approach proposed by Eisenhardt (1989).

The remainder of this research paper is organised as follows. Section 2.3.4 highlights the literature related to data governance to aid understanding of its concepts. Section 2.3.5 outlines the research methodology and explains the research approach and data gathering and analysis techniques used. Section 2.3.6 presents the CSFs identified for data governance, together with detailed descriptions, followed by a section on the possible interconnectedness of the CSFs. Finally, section 2.3.8 presents the concluding remarks and areas of further research.

2.3.4 Data Governance Background

Data governance has received much attention in both academic and practitioner communities. The concept has been developed over the last ten years whereby data are considered as valuable assets and as a strategic function within the organisation's structure and are thus placed under corporate governance (Vayghan, Garfinkle, Walenta, Healy, & Valentin, 2007; Wende, 2007). Data governance focuses on who holds the decision rights related to data assets in an organisation (Khatri & Brown, 2010; Otto, 2011) in order to ensure the quality, consistency, usability, security, privacy, and availability of the data (Cohen, 2006; Panian, 2010).

Rau (2004, p. 35) refers to governance as "*the way the organization goes about ensuring that strategies are set, monitored, and achieved*". Horne (1995) connected governance with the optimal use of assets and outlined how data as an asset drives the importance of the governance of data within an organisation. The concept of data as an asset emerged with a report by the Hawley Committee in 1994, which defined data assets as "*data that is or should be documented and that has value or*

potential value" (Oppenheim, Stenson, & Wilson, 2003, p. 159). Therefore, the main driver of data governance is the consideration data as an asset in an organisation (Panian, 2010).

Several data governance models have been proposed which enable us to understand the boundaries of data governance and related functions (Cheong & Chang, 2007; Guetat & Dakhli, 2015; Khatri & Brown, 2010; Lajara & Maçada, 2013; Otto, 2011; Vayghan et al., 2007; Wende & Otto, 2007).

In our recent study (Chapter 2, paper 1), we analysed academic and practitioner publications on data governance and proposed a universal data governance activities model. We identified a set of data governance activities that interconnect three main constructs: 1) action (plus) 2) area of governance (plus) 3) decision domain. These activities can be seen in Figure 2-10.

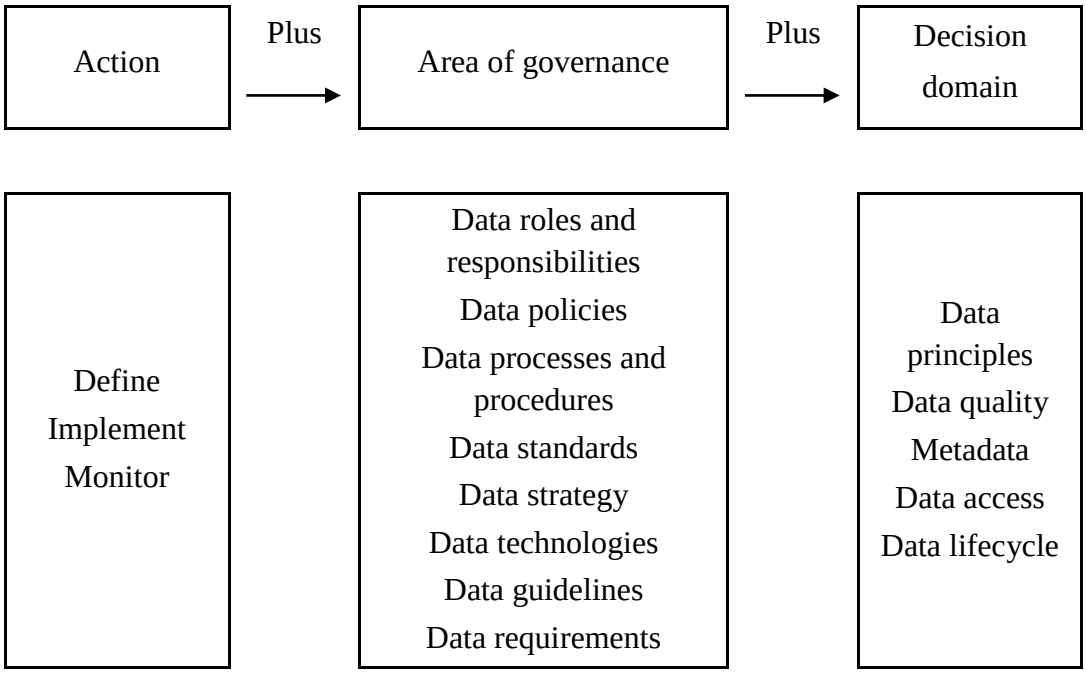

Figure 2-10 Illustrates the three data governance activities' constructs, including possible values (Chapter 2, paper 1).

From the above model, it can be seen that there are eight areas of governance reported in the literature across the five decision domains suggested by Khatri and Brown (2010) for which the data governance function should contain decisions around defining, implementing, and monitoring in order to ensure a successful data governance programme. Our analysis shows a lack of research around data governance, particularly in the implementation and monitoring actions. There is more focus in the literature on the defining action, which indicates a somewhat embryonic understanding of data governance.

However, the activities are considered universal data governance activities and can vary from one organisation to another. These can be a guide to understanding data governance-related concepts and boundaries, as they are a collection of the activities reported in different publications. As a result, we needed to understand how governance is actually executed within practice. Therefore, we decided to follow an indicative approach to build theory from a case study. In the next section, we provide a detailed description of our research approach to building theory.

2.3.5 Research methodology

The theory building research strategy proposed by Eisenhardt (1989) provides a clear process for conducting research that aims to build theories from one or more case studies. According to Eisenhardt (1989), the main driver for building theory from a case study is when little is known about a phenomenon and, therefore, the process does not rely on previous literature or prior empirical evidence. Hence, interpretive qualitative research is an appropriate research design to apply when exploring CSFs by conducting a case study (Koh et al., 2011). Many scholars have

investigated and explored CSFs in certain IS domains and applied qualitative methods using either single case or multiple case study designs (e.g., Butler & Fitzgerald, 1999; Guynes & Vanecek, 1996; Sammon & Adam, 2008).

2.3.5.1 Case Background

Al Rajhi Bank was founded in 1957 and is considered to be one of the largest Islamic banks in the world, with total assets of US$80 billion, paid-up capital of US$4.33 billion and an employee base of over 9,600 associates. With over 58 years of experience in banking and trading activities, the various individual establishments under the Al Rajhi name were merged under the umbrella of 'Al Rajhi Trading and Exchange Corporation' in 1978 and it was in 1988 that the bank was established as a Saudi share-holding company. With an established base in Riyadh, Saudi Arabia, Al Rajhi Bank has a vast network of over 500 branches, over 118 dedicated women's branches, more than 4,100 ATMs, 46,000 point-of-sale terminals installed with merchants, and the largest customer base of any bank in the Kingdom of Saudi Arabia, in addition to 170 remittance centres across the kingdom.

Al Rajhi Bank recorded net income profits of US$2,166 million in 2016. The bank operates in multiple segments and continues to grow through the diversification of income resources and development of the investment and corporate banking sectors, which are built on a strong retail banking base.

Internationally, Al Rajhi Bank currently has 24 branches in Malaysia. It also started operations in Kuwait in 2010 with a fully-fledged branch offering retail and corporate banking solutions. In addition, the bank started activities in Jordan in

2011, offering its customers innovative and comprehensive banking products and services to help them enhance their lifestyles. The bank now has six branches in Jordan, through which it serves the top three occupied districts in Jordan.

This case was selected because Al Rajhi bank deals with massive amounts of data which is distributed between different systems. Also, in the banking industry, data governance is considered to be a vital function in the organisation as they deal with financial data as well as sensitive customer data. Within Al Rajhi bank, some stories were considered to be successes and others as failures, which can be valuable to our research.

2.3.5.2 Data gathering

The CSF approach was introduced by Rockart (1979), who defined CSFs as the "*areas of activity that should receive constant and careful attention from management*" (p. 85). The CSFs approach has been widely investigated and used in information systems (IS) research and in practice over the last three decades (Shah et al., 2007; Tan et al., 2009) and remains a valid research method for making sense of a problem by identifying potential factors that influence a community of practice (Caralli et al., 2004; Lam, 2005).

Interviews are considered the most appropriate data gathering technique for collecting rich and detailed data from industry experts (Koh et al., 2011). Interviews are subject to the amount of control utilised by the researcher during the interview and the degree of structure required (Esterberg, 2002). In terms of identifying CSFs, Rockart (1979) suggests conducting separate interviews with executives

individually. The interviews conducted for this research were aimed at identifying the business goals that indicate CSFs. Therefore, this research employed semi-structured interviews, which enabled the researchers to explore the CSFs for data governance. We developed a data collection procedure based on the CSF approach in Rockart (1979) (Figure 2-11).

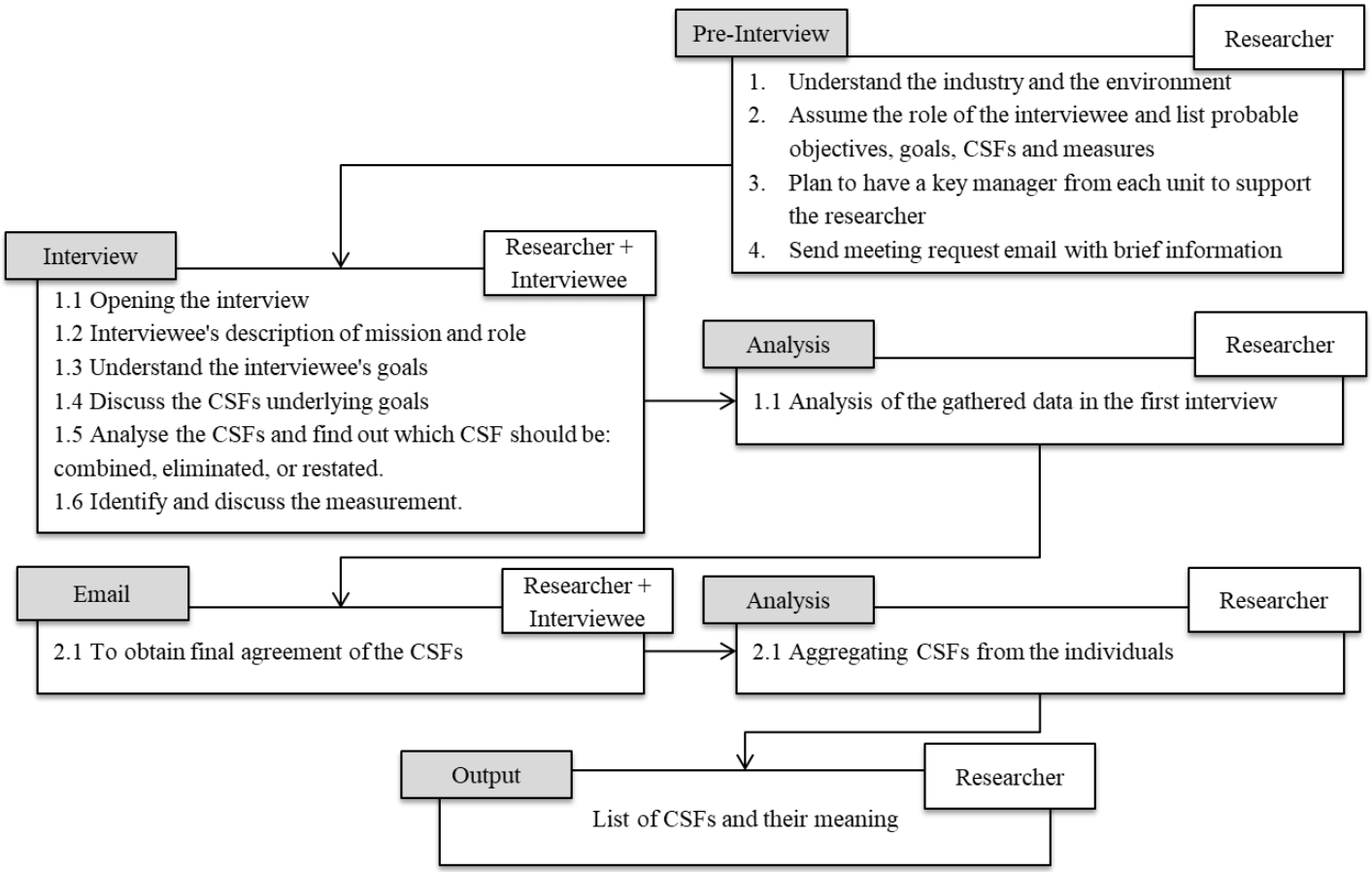

Figure 2-11 Data gathering approach.

Semi-structured interviews were the method used for data gathering. Fifteen individual interviews were conducted at Al Rajhi Bank with personnel at the managerial levels of both business and IT departments (see Table 2-15 for a list of the interviewees' positions and the duration of the interviews). These interviews were conducted in two different periods. The researchers decided to stop interviewing more people at the point at which information started to be repeated and the material collected was sufficiently rich to cover the majority of the data governance aspects.

All the interviews were started with an introduction of the research objective. Each interviewee was then asked to begin talking about the data-related activities in his/her department. Then, during the interviews, we identify the related CSFs for data governance. In many cases, the interviewer explained the data governance programme from the perspective of the five decision domains to make sure the interviewee understood the meaning of data governance. During the interviews, the interviewer attempted to keep the discussion to data-governance-related topics in order to concentrate the interviews around the research subject.

Some of the interviews were conducted in Arabic and others in English, depending on the English-language level of the interviewee. All the interviews were transcribed word-by-word and those conducted in Arabic were translated into English by a third party in order to avoid bias. The transcripts were then reviewed with the recording in order to supply any missing words. Finally, due to the transcripts having been translated, they were reviewed to ensure that they were true to the meaning of the original interview.

Table 2-15 List of interviewees' positions and related section (IT or business) and interview duration.

Position	IT/Business	Interview duration (minutes)
Product Manager for Mobile Banking	Business	60
Head of Remittance	Business	70
Head of Government Relations Department	Business	30
Head of Alternative Channels	Business	50
Head of Call Centre	Business	60
Head of Data Cleansing Project	Business	70
Head of Internet and Mobile Banking	Business	60
Product Manager	Business	40
Head of MIS & HR Payments	Business	70
Head of Risk Systems & Data Governance	Business	70
(1) Senior Systems Analyst (Data Warehouse)	IT	40
(2) Senior Systems Analyst (Data Warehouse)	IT	40
Senior Systems Analyst (Internet Banking)	IT	60
Oracle Analyst	IT	60
Head of IT Risk	IT	30

2.3.5.3 Data analysis

Qualitative data analysis is not well formulated (Miles, 1979), and there are probably as many approaches as there are researchers (Eisenhardt, 1989). In addition, the emphasis of qualitative data analysis is on "*sense making*"

(Bhattacherjee, 2012, p. 113), so a coding technique by Strauss and Corbin (1990) was adopted in this research in a way that serves the purpose of the research objective. Coding is one of the techniques widely used in analysing qualitative data in order to build theory from a case study (Buchwald et al., 2014; Tallon et al., 2013; Tan et al., 2015). In the following data analysis, there are, as outlined by Strauss and Corbin (1990), three types of coding: open, axial, and selective (see Table 2-16). These coding techniques aim to generate concepts from field data (Walsham, 2006). According to Strauss and Corbin (1990, p. 57), coding "*represents the operations by which data are broken down, conceptualized, and put back together in new ways*".

Table 2-16 Open, axial, and selective coding definitions by Strauss and Corbin (1990).

Coding technique	**Definition**
Open coding	"*The process of breaking down, examining, comparing, conceptualizing, and categorizing data*" (p. 61).
Axial coding	"*A set of procedures whereby data are put back together in new ways after open coding, by making connections between categories. This is done by utilizing a coding paradigm involving conditions, context, action/interactional strategies and consequence*" (p. 96).
Selective coding	"*The process of selecting the core category, systematically relating it to other categories, validating those relationships, and filling in categories that need further refinement and development*" (p. 116).

Open coding is a process that aims to identify the concepts or key ideas that are hidden within data that are likely to be related to the phenomenon of interest

(Bhattacherjee, 2012). Concepts and categories are generated in the open coding stage (Glaser, 1992). Strauss and Corbin (1990) state that the concepts that appear to be similar are grouped together under a higher-order, more abstract concept called a category.

The second reading of the data is considered during axial coding (Dezdar & Sulaiman, 2009), which is performed simultaneously with open coding (Bhattacherjee, 2012; Strauss & Corbin, 1990). During this stage, the categories are refined in order to be linked in the form of relationships. Strauss and Corbin (1990) suggest that, in order to identify the relationship between data, a paradigm model should be used that consists of causal conditions, the phenomenon, the context, intervening conditions, action/interaction strategies, and consequences. Using this model enables the researcher to think systematically about the data in order to relate them (Strauss & Corbin, 1990).

Finally, selective coding begins when researchers identify a potential core category (Tan et al., 2015), focusing then on the core categories and related categories that accrued in the axial coding. This involves comparing the core categories with the raw data by telling the story of the core categories that emerge (Strauss & Corbin, 1990).

For this research, after preparing all the interview transcripts, the data analysis was commenced by reading each transcript sentence by sentence and following an open coding technique. After coding the first two interviews, axial coding was commenced in an iterative manner as categories started to emerge (see Figure 2-12). The five decision domains identified by Khatri and Brown (2010), were used

to break down the phenomenon into paradigm models (see Figure 2-13: Paradigm model constructs) in order to clarify the relationships between the categories that emerged during the open coding analysis. Therefore, the axial coding procedure resulted in five paradigm models that identify the relationships between the categories. The researchers were then able to identify selective coding for the core categories and validate the concepts that emerged in an iterative manner. The core categories are considered later as CSFs for data governance.

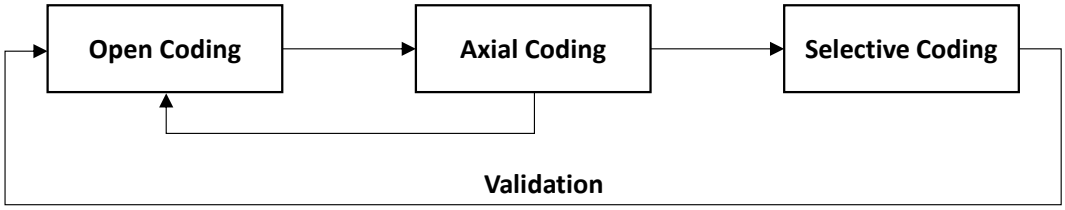

Figure 2-12 Open, axial, and selective coding iterative process.

The five decision domains identified by Khatri and Brown (2010), namely, 1) data principles, 2) data quality, 3) metadata, 4) data access, and 5) data life cycle (see Table 2-17) are used as an initial lens to identify the CSFs for data governance.

Table 2-17 Decision domains for data governance (Khatri & Brown, 2010).

<table>
<tr><td colspan="3">Data principles “Clarifying the role of data as an asset”</td></tr>
<tr><td rowspan="2">Data quality
“Establishing the requirements of intended use of data”</td><td>Metadata
“Establishing the semantics or “content” of data so that it is interpretable by the users”</td><td rowspan="2">Data life cycle
“Determining the definition, production, retention and retirement of data”</td></tr>
<tr><td>Data access
“Specifying access requirements of data”</td></tr>
</table>

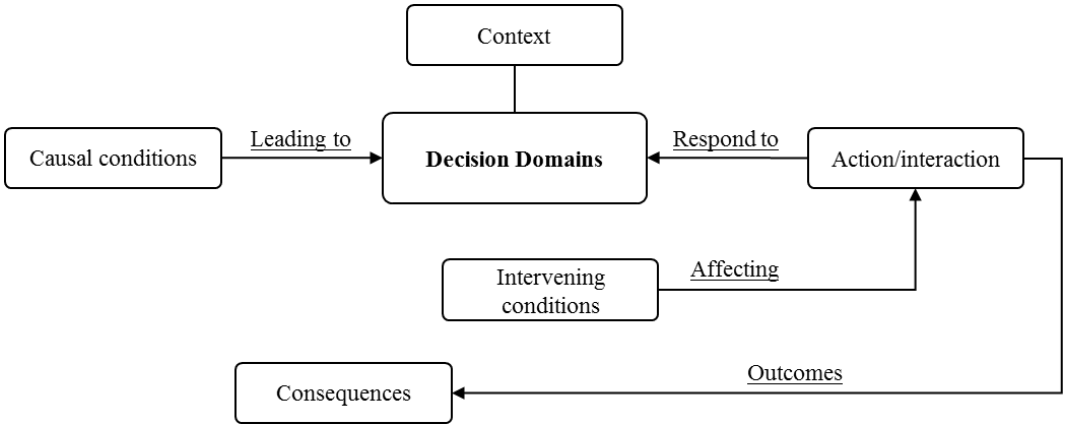

Figure 2-13 Paradigm model constructs.

Some of the categories that emerged are associated with more than one decision domain, due to the context of the original concepts. For example, the category 'Employee awareness' was associated as a causal condition of 'Data principles' where one of the interviewees stated the following in a general comment about data governance: "*Our people here are well educated, but do they have the concepts of how to work on data governance? No, they don't*". In contrast, in another interview, the category 'Employee awareness' was associated with 'Data quality' when the interviewee stated in the context of the data quality level that "*there was no awareness and 90% of the problem is that the employee doesn't have awareness*".

From coding the 15 interviews, it was found that the majority of the concepts were associated with 'Data quality' as a decision domain for data governance, as can be seen in Figure 2-14. This is not surprising, as data quality is considered a fundamental element of a data governance programme. This is followed by 'Data principles', whereby the strategic initiatives are associated with the overall data governance programme.

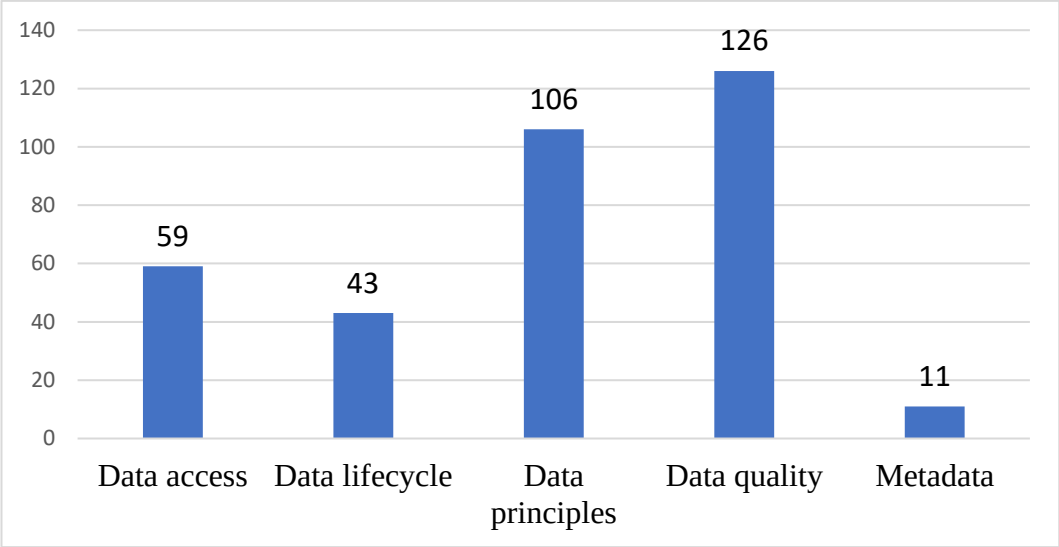

Figure 2-14 Frequency count of the categories associated with each of the five decision domains.

The coding procedure for the 15 interviews resulted in 345 concepts that related to data governance. The 345 concepts generated 89 categories. Using the paradigm models, the researchers identified the relationships between the 89 categories, which enabled the creation of seven core categories during the selective coding phase. Figure 2-15 illustrates the processes together with examples of concepts, category relationships, and the core category, namely, 'employee data competencies' and its cause and action/interaction.

Open coding	**Concept #1:** Many of the employees don't have a good understanding of the English language	**Excerpt:** *For example, the English name field is left to the clerk's sensibilities in terms of the spelling, and many of them don't know English very well.*
	Concept #2: Certain competencies	**Excerpt:** *The clerk has certain competencies.*
	Concept #3: Teaching the policy and procedure to employees	**Excerpt:** *So, as I said, it should be solved by teaching the procedures and policies to the employees.*
	Concept #4: Educate business people in data governance	**Excerpt:** *Hence, the business people should do the governance of the data and they have to be educated about it.*

Axial coding	**Category #1:** Employee competency levels	**Concept #1:** Many of the employees don't have a good understanding of the English language. **Concept #2:** Certain competencies.	**Paradigm model** Employee competency level is a causal condition for data quality.
	Category #2: Employee training	**Concept #3:** Teaching the policy and procedure to employees. **Concepts #4:** Educate the business people on data governance.	**Paradigm model** Training is an action/interaction for data quality.

Selective coding	**Core category #1:** Employee data competencies	**Cause:** Employee competency level
		Actions: Employee training

Figure 2-15 Example of the data coding procedure for the 'Employee data competencies' CSF.

2.3.6 Findings

2.3.6.1 CSFs for data governance

This subsection discusses the CSFs identified as a result of the five paradigm models and the selective coding findings. Seven core categories emerged that are considered to be CSFs for data governance: 1) Employee data competencies, 2) Clear data processes and procedures, 3) Flexible data tools and technologies, 4) Standardised easy-to-follow data policies, 5) Established data roles and responsibilities, 6) Clear inclusive data requirements, 7) Focused and tangible data strategies.

Table 2-18 shows the seven CSFs for data governance, ranked based on the frequency count of the concepts they reflect. The CSFs are associated with the most obvious causes and subsequent actions/interactions. Causes are positive or negative things that specify possible to consider CSFs, whereas the actions/interactions are the things that are recommended to be performed in order to address the CSFs. These causes and actions/interactions are abstracted from the original results of the axial coding. In the next section, each of the seven CSFs is described in greater detail in order to clarify the meaning and boundaries of each stated CSF.

Table 2-18 CSFs for data governance associated with main causes and actions/interactions

Rank	Critical success factor	Cause	Action/interaction
1	Employee data competencies	Questionable employee competency level and top management awareness.	Increase employee awareness and training.
2	Clear data processes and procedures	Significant manual data entry.	Have appropriate data processes and procedures and embed them into the systems.
3	Flexible data tools and technologies	Data integration and ability to embed data policies, processes, and procedures.	Have appropriate IT infrastructure and integrated data.
4	Standardised easy-to-follow data policies	Lack of clear data policies	Embed data policies into the systems.
5	Established data roles and responsibilities	Unclear roles and responsibilities.	Assign a committee for data governance and define the data owners.
6	Clear inclusive data requirements	Understanding of data requirements and communication issues.	Have the right data requirements and comply with regulations.
7	Focused and tangible data strategies	Understanding the importance of the data.	Consider data as a strategic element and management reinforcement of this ethos.

CSF #1: Employee data competencies

The employee data competencies CSF covers data governance activities that involve human action and is, based on our case analysis, the highest-ranked CSF for data governance. Employee data competencies directly impact the defining, implementing, and monitoring of data processes and procedures, as well as data policies and data requirements. Thus, it determines an employee's ability to handle these data governance activities.

The competencies of all employees, from senior executives to entry-level workers, are important due to their involvement in various data governance activities at various points in time. For example, establishing an overall data governance strategy requires certain top managers to have certain competencies. Based on our analysis, such competencies would be needed to treat data as a strategic asset. Furthermore, dealing with data entry and access also requires employees to have a minimum set of capabilities and a certain level of awareness with regard to handling the organisation's data. For example, the practice of manual data entry depends on a certain level of employee capabilities and awareness of data processes and procedures. In addition, due to the sensitivity of banking data and privacy requirements, our analysis highlighted the importance of increasing employee awareness around viewing and modifying the data from both business and IT departments in order to avoid the misuse of customers' information.

Different actions/interactions are recommended in order to ensure appropriate employee data competencies. The most important action/interaction is 'training', such as continuous training in dealing with and implementing data policies as well

as data processes and procedures, and includes internal and external training. For example, the bank addresses data entry problems by building up a team responsible for training the operations and branch managers, as stated by one of the executive managers of a data cleansing project: "*We have built a training team from the CIF Department who are specialists in this field. They trained the operations managers and branch managers*". It is also vital to increase employees' awareness of the criticality of data in terms of entering the right information, as well as when accessing sensitive material.

CSF #2: Clear data processes and procedures

Based on our analysis, clear data processes and procedures ranked second in importance as a CSF for data governance. This is not surprising, as the bank generally has data policies in place which should be detailed and operationalised during activities related to data processes and procedures. Therefore, clear data processes and procedures are evenly coded among the five decision domains, apart from metadata. This includes all the detailed activities related to data flow, data integration, data authorisation processes, data validation, and more.

The absence of data processes and procedures resulted in doubts relating to trusting the data. This is due to different reasons, one of which is knowing that there are no clear data processes and procedures, as stated by one interviewee: "*When a person gets the data, they assume the wrong intention, and the reason for that is that there are no clear procedures for the data*". Another reason is missing a part of the data processes and procedures, such as data testing. For example, when establishing a new product in the bank, the resulting data are not tested enough to have the proper

report, as stated by the department responsible for evaluating the result of each product in different branches: "*our problem is that we start to use data before we run the right test on it*".

Based on our analysis, different actions/interactions are recommended in order to achieve effective data processes and procedures in the bank. At the top of these actions/interactions is 'Embedding the data processes and procedures into the system', such as when an interviewee was talking about the processes of attaching the right documents: "*That would be facilitating the processes itself, when you scan or photocopy the customer's ID and then attach it to his file, it's expected that the fields become embedded in the system*". This includes considering mandatory fields, validation methods, and other data flow requirements. Finally, the current processes and procedures need to be re-checked and updated, as stated by the Head of Alternative Channels: "*Also, you will find the manual data processes require re-checking*".

CSF #3: Flexible data tools and technologies

Flexible data tools and technologies consist of all the activities related to dealing with software and hardware that affect the data in an organisation, including the presentation and storage of data. Flexible data tools and technologies were, according to our analysis, coded among all five decision domains. However, the majority of the codes were associated with the data life cycle, as it involves decisions relevant to the operationalising and processing of the data throughout different systems.

Our analysis shows a significant impact of flexible data tools and technologies on other CSFs, such as data policies, as well as clear data processes and procedures in terms of embedding them into the right system with the correct format. This includes, for example, making some fields mandatory or having an automated validation method. The tools and technologies also involve the implementation of data requirements. Therefore, having strong data tools and technologies enables the success of other CSFs.

Having appropriate IT infrastructure and integrated data is recommended to address flexible data tools and technologies. This includes setting up advanced technologies that enable data integration in order to automate the validation of the data. As stated by one interviewee: "*Frankly, the solution is that we get more automated tools to capture this data. That way, it's automated or verified from a reliable source*". In addition, it is recommended that systems should be thoroughly tested through a testing procedure, as well as flexible enough to incorporate future changes. Finally, it is important to take into account the privacy and availability of the data while integrating internal and external systems.

CSF #4: Standardised easy-to-follow data policies

Standardised easy-to-follow data policies play a fundamental role as a CSF for data governance. Data policies are short statements that define the high-level guidelines and rules necessary for dealing with data. In the context of a CSF, data policies should include definition, implementation and monitoring. Our analysis shows that standardised easy-to-follow data policies are associated with all five decision domains, particularly for data principles. Furthermore, data access is reliant on data

policies existing especially for banking data, since banks deal with critical data that requires a high degree of privacy.

Based on our analysis, the absence of data policies for certain data made employees uncomfortable about making decisions that relied on these data due to their not knowing how the data were processed. Accessing unneeded data affecting privacy might also have an effect on business performance. Certain other characteristics of having data policies were also suggested by our analysis. For example, data policy documents should follow a certain template in order to be understood by all the employees who deal with them, as well as keeping the policy statements basic, simple and up-to-date to ensure that employees appreciate the value of following the guidelines.

Several actions/interactions are recommended in order to have effective data policies in place, which include having a strong, clear, simple, and easy-to-follow data policy. However, having a defined data policy is not enough to achieve successful data governance. Formulating implementation methods and having them in place is highly recommended. For example, as drawn from our analysis, embedding data policies into a system is strongly recommended and involves having mandatory fields for data entry and a validation method especially for sensitive data, such as, in our case, recording the mobile number of the customer whereby the customer receives encrypted data via text message. Monitoring policies and updating them is another of the essential activities related to this CSF that are highly recommended following our analysis. For example, in the context

of data policies, the Head of Internet Banking stated: "*after that, what you need to do is to do a periodic audit on it*".

CSF #5: Established data roles and responsibilities

Established data roles and responsibilities should also be considered in the context of data governance. It is important to identify the individual(s) responsible for the data-related activities in the organisation, such as who defines the policies and processes for the data as well as assigning the duties for the actions related to data. In addition, from a strategic point of view, established data roles and responsibilities include the data governance function in the organisation. Therefore, established data roles and responsibilities categories are mainly coded under data principles as the decision domain.

Our analysis shows that the employees in the bank experienced imprecise roles and responsibilities that were caused either by the roles being unclear or having unclear assignments. This confirms established data roles and responsibilities as a CSF for data governance. For example, having good processes in place without clear roles and responsibilities leads to mistakes in dealing with data, as stated by one of the interviewees: "*This dual control process is implemented as written in the procedures, but who is the monitor? Who edits these fields? Mistakes happen unfortunately on an ad hoc basis*". Therefore, unclear data roles and responsibilities has a negative effect on the success of data governance.

Different actions/interactions are recommended from our analysis, such as setting up a committee for data governance as well as identifying data owners. For example, it was stated by an interviewee that "*there should be a committee so you*

can formulate this". It is recommended, therefore, to assign a committee to deal with all data governance activities. The Head of the Call Centre also remarked: "*So, we are the decision-maker in the call centre, what data can be shown? But at the bank level, no, there are different departments for different systems*". This indicates the need for a clear definition of data owners.

CSF #6: Clear inclusive data requirements

In the context of data governance, data requirements are the needs that are initially requested by business of IT with regard to data. Clear inclusive data requirements define all aspects of data implementation, such as data flow and integration. Our analysis shows that the business owner should initially understand the data requirements and then communicate properly with IT in order to explain the data needs clearly.

The largest element of data requirements in our analysis are those associated with the data regulations that come from either external regulators, such as the central bank or other corresponding banks; for example, one interviewee stated that "*we got regulations from the central bank enforcing us to make the changes*", or internal regulations, such as the Compliance Department at the bank. These regulations are considered the main part of the data requirements for data governance due to the banking industry being well regulated. In addition, the clear inclusive data requirements include how data are used for different types of reports. As stated by one of the IT developers, this means "*how they want their data to be presented and calculated in their way*". Therefore, based on our analysis, the IT Department in the bank suffers from a lack of clear data requirements from the business owner of the

data, as can be seen from a comment by one of the developers: "*I only add what I expect I will need in the reports, and what other departments and authorities may need, in a way that I think is correct, unless I have clear requirements*".

There is a clear relationship between clear inclusive data requirements and employee data competencies, as understanding the right requirements is considered the main action/interaction to address data requirements. This is heavily based on employee data competencies. The data owners should also communicate with the implementers in a formal and detailed method for each and every data requirement, including the data flow as well as mandatory fields and validation methods.

CSF #7: Focused and tangible data strategies

Focused and tangible data strategies include planning for data governance in order to achieve its goals, as well as the main activities related to considering data as assets. Furthermore, the short and long term objectives that relate to data governance are included. Therefore, based on our analysis, data strategy categories are mainly related to data principles as a decision domain.

Understanding the importance of the data and considering them as assets confirms focused and tangible data strategies as a CSF for data governance. For example, one of the team members in the data cleansing project stated that "*data cleansing is one of the projects considered strategic in the bank*". Our analysis also shows that poor planning for the future negatively impacts on data when focused and tangible data strategies are absent, as stated in this excerpt: "*Also, there was poor planning for the future needs, we are talking about 20 years ago*". However, based

on our analysis, the bank recently started to consider data as strategic elements, which has an influence on the success of data governance.

In contrast, based on our analysis, considering data as strategic elements is the main recommended action for addressing focused and tangible data strategies. Top management enforcement should also be taken as an action, as stated by the Head of IT Risk: "*you need enforcement from top to bottom*". This includes considering the assignment of a top committee for data governance.

2.3.7 Possible interconnectedness of CSFs

As an outcome of our analysis, the possible interconnectedness of the CSFs is also explored. This enables a better understanding of the multiplicity of effects of each CSF. The interconnections articulate the existing impact of the absence/presence of one of the CSFs on each other. Table 2-19 presents the possible interconnectedness between the CSFs.

For example, 'employee data competencies' defines the competency requirements for each role, which impacts the CSF relating to 'established data roles and responsibilities'. Therefore, a failure to have employee data competencies may cause a failure in establishing data roles and responsibilities. On the other hand, 'established data roles and responsibilities' directly impact all the other CSFs, as all the other CSFs are performed within the data governance structure in which organised roles and responsibilities are required. Hence, a failure to establish data roles and responsibilities directly impacts the performance of the other CSFs.

Table 2-19 Possible interconnectedness of CSFs.

CSF	Has an impact on	Interconnection
Employee data competencies	Established data roles and responsibilities	Defines the competency requirements for each role
Flexible data tools and technologies	Clear data processes and procedures	Embeds the data processes and procedures into the systems
	Standardised easy-to-follow data policies	Embeds the data policies into the systems
Established data roles and responsibilities	All CSFs	Defines who does what and assigns responsibilities for each activity
Clear inclusive data requirements	Flexible data tools and technologies	Drives the required tools and technologies for data governance

The absence of one of the CSFs not having a direct impact on other CSFs, such as 'standardised easy-to-follow data policies', does not mean that the CSF is of lower importance compared with the others. The potential interconnectedness only shows the possible impact of the presence or absence of one CSF on another.

2.3.8 Concluding Remarks and Further Research

This research has attempted to contribute to the existing knowledge of data governance by addressing the CSFs for data governance. This research followed the approach of theory building by conducting several semi-structured interviews within a single case study organisation. The interviews were transcribed and

prepared for analysis by applying open, axial, and selective coding. Seven CSFs for data governance have been identified based on our analysis of the case data. These CSFs are associated with recommended actions/interactions in order to enable organisations to have successful data governance. Possible interconnectedness between the CSFs found has also been presented.

We found that establishing a successful data governance programme requires a high level of attention from the stakeholders with regard to employee competencies, as well as a need to start by establishing data roles and responsibilities. These two CSFs enable the activities around the other CSFs to be established successfully and reflect the value of having a successful data governance programme.

Certain limitations in this research could be addressed in a future study. One of the limitations is that this research was based on a single case study, which can only show part of the picture. Conducting another case study in a different industry is recommended, in order to have a universal model of CSFs for data governance. It is also recommended that further research examine the interconnectedness between CSFs in order to better understand the implications of the presence or absence of each factor.

2.3.9 References

Bhattacherjee, A. (2012). Social science research: principles, methods, and practices (2nd ed.). Tampa, FL: CreateSpace Independent Publishing Platform.

Buchwald, A., Urbach, N., & Ahlemann, F. (2014). Business value through controlled IT: toward an integrated model of IT governance success and its impact. Journal of Information Technology, 29(2), 128-147. Retrieved from https://link.springer.com/article/10.1057/jit.2014.3

Butler, T., & Fitzgerald, B. (1999). Unpacking the systems development process: an empirical application of the CSF concept in a research context. The Journal of Strategic Information Systems, 8(4), 351-371. doi:10.1016/S0963-8687(00)00027-5

Caralli, R.A., Stevens, J.F., Willke, B.J., & Wilson, W.R. (2004). The critical success factor method: establishing a foundation for enterprise security management. Retrieved from http://www.dtic.mil/get-tr-doc/pdf?AD=ADA443742

Cheong, LK & Chang, V 2007, 'The need for data governance: A case study', Proceedings of the 18th Australasian conference on information systems, ACIS 2007 Proceedings, pp. 999-1008.

Cohen, R. (2006). What's in a name? Data governance roles, responsibilities and results factors. DM Review, 8.

Dezdar, S., & Sulaiman, A. (2009). Successful enterprise resource planning implementation: taxonomy of critical factors. Industrial Management & Data Systems, 109(8), 1037-1052. doi:10.1108/0263557091099

Eisenhardt, K.M. (1989). Building Theories from Case Study Research. The Academy of Management Review, 14(4), 532-550. doi:10.5465/AMR.1989.4308385

Esterberg, K.G. (2002). Qualitative methods in social research. Boston, MA: McGraw Hill.

Glaser, B. G. (1992). Emergence vs forcing: Basics of Grounded Theory Analysis. Mill Valley, CA: Sociology Press.

Guetat, SBA & Dakhli, SBD 2015, 'The architecture facet of information governance: The case of urbanized information systems', Procedia Computer Science, vol. 64, pp. 1088-1098. Available from: http://0-dx.doi.org.library.ucc.ie/10.1016/j.procs.2015.08.564.

Guynes, C.S., & Vanecek, M.T. (1996). Critical success factors in data management. Information & Management, 30(4), 201-209. doi:10.1016/0378-7206(95)00053-4

Holt, V., Ramage, M., Kear, K., & Heap, N. (2015). The usage of best practices and procedures in the database community. Information Systems, 49, 163-181.

Horne, N.W. (1995). Information as an asset—The board agenda. Computer Audit Update, 1995(9), 5-11. doi:10.1016/0960-2593(95)90246-5

Information Builders. (2013), Driving Better Business Performance With a

Practical Data Strategy, [White paper].

Khatri, V., & Brown, C.V. (2010). Designing data governance. Communications of the ACM, 53(1), 148-152. doi:10.1145/1629175.1629210

Koh, S.C.L., Gunasekaran, A., & Goodman, T. (2011). Drivers, barriers and critical success factors for ERPII implementation in supply chains: A critical analysis. The Journal of Strategic Information Systems, 20(4), 385-402. doi: 10.1016/j.jsis.2011.07.001

Lam, W. (2005). Investigating success factors in enterprise application integration: a case-driven analysis. European Journal of Information Systems, 14(2), 175-187. Retrieved from https://link.springer.com/article/10.1057/palgrave.ejis.3000530

Lajara, TT & Maçada, ACG 2013, 'Information governance framework: The defense manufacturing case study', Proceedings of the nineteenth Americas conference on information systems. Available from: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1704&context=amcis2013. [15-17 August 2013].

Miles, M.B. (1979). Qualitative Data as an Attractive Nuisance: The Problem of Analysis. Administrative Science Quarterly, 24(4), 590-601. Retrieved from http://www.jstor.org/stable/2392365

Oppenheim, C., Stenson, J., & Wilson, R.M.S. (2003). Studies on Information as an Asset I: Definitions. Journal of Information Science, 29(3), 159-166. doi:10.1177/01655515030293003

Otto, B. (2011). Organizing data governance: findings from the telecommunications industry and consequences for large service providers. Retrieved from http://aisel.aisnet.org/cgi/viewcontent.cgi?article=3610&context=cais

Otto, B. (2015). Quality and value of the data resource in large enterprises. Information Systems Management, 32(3), 234-251. doi:10.1080/10580530.2015.1044344

Panian, Z. (2010). Some practical experiences in data governance. World Academy of Science, Engineering and Technology Management, 62, 468-475.

Rau, K. G. (2004). Effective governance of it: Design objectives, roles, and relationships. Information Systems Management, 21(4), 35-42. doi:10.1201/1078/44705.21.4.20040901/84185.4

Redman, T. C. (2013). Data's credibility problem. Harvard Business Review, 91(12), 84-88.

Rockart, J.F. (1979). Chief executives define their own data needs. Harvard business review, 57(2), 81-93. Retrieved from http://europepmc.org/abstract/med/10297607.

Sammon, D., & Adam, F. (2008). Proceedings from 2008 International Conference on Information Systems (ICIS): Justifying an ERP investment: critical success factors for transformation investments. Paris, France.

Shah, M.H., Braganza, A., & Morabito, V. (2007). A survey of critical success factors in e-Banking: an organisational perspective. European Journal of Information Systems, 16(4), 511-524. Retrieved from https://link.springer.com/article/10.1057/palgrave.ejis.3000693

Strauss, A., & Corbin, J. (1990). Basics of qualitative research: Grounded theory procedures and techniques. Thousand Oaks, CA: SAGE Publications, Inc.

Tallon, P.P., Ramirez, R.V., & Short, J.E. (2013). The Information Artifact in IT Governance: Toward a Theory of Information Governance. Journal of Management Information Systems, 30(3), 141-178. doi:10.2753/MIS0742-1222300306

Tan, B., Pan, S.L., Lu, X., & Huang, L. (2015). The Role of IS Capabilities in the Development of Multi-Sided Platforms: The Digital Ecosystem Strategy of Alibaba. com. Journal of the Association for Information Systems, 16(4), 248-280.

Tan, W.G., Catersteel, A., & Toleman, M. (2009). Implementing it service management: A case study focussing on critical success factors. Journal of Computer Information Systems, 50(2),1-12. doi:10.1080/08874417.2009.11645379

Walsham, G. (2006). Doing interpretive research. European journal of information systems, 15(3), 320-330. Retrieved from https://link.springer.com/article/10.1057/palgrave.ejis.3000589

Weber, K., Otto, B., & Österle, H. (2009). One Size Does Not Fit All--A Contingency Approach to Data Governance. Journal of Data and Information Quality (JDIQ), 1(1). doi:10.1145/1515693.1515696

Wende, K. (2007). Proceedings from 18th Australasian Conference on Information Systems (ACIS): A model for data governance–Organising accountabilities for data quality management. Toowoomba, Australia: The University of Southern Queensland.

Wende, K & Otto, B 2007, 'A contingency approach to data governance'. In: 12th International Conference on Information Quality (ICIQ-07), pp. 163-176. Available from: https://www.alexandria.unisg.ch/213308/. [10 November 2007].

Vayghan, JA, Garfinkle, SM, Walenta, C, Healy, DC, & Valentin, Z 2007, 'The internal information transformation of IBM', IBM Systems Journal, vol. 46, no. 4, pp. 669-683.

2.4 Paper 4

Critical success factors for data governance in the telecommunications industry

2.4.1 Abstract

The objective of this research study is to present the critical success factors (CSFs) for data governance. This paper reports on a single case study in a large telecommunications service provider in Saudi Arabia. The data are gathered through semi-structured interviews that follow the CSF approach and analysed by applying open, axial, and selective coding techniques. The findings of this research are presented as nine CSFs, which are ranked in order of importance. Based on our case analysis, having 'proper data integration strategies' was the most important factor for data governance.

2.4.2 Keywords:

Data governance, CSFs, case study, open coding, axial coding, selective coding

2.4.3 Introduction

The ability to govern data is playing an increasingly critical role in big organisations in order for them to drive business success. A recent study by Nagle and Sammon (2017) shows that data governance is, however, a problem area for the majority of organisations. The successful implementation of a data governance programme would give an organisation a competitive advantage, as well as protecting it from incidents that could affect the entity (Hassan & Chindamo, 2017). Data governance has the aim of considering data as strategic assets and driving the value of the business.

Although data governance is considered a relatively new and emerging subject (Kamioka, Luo, & Tapanainen, 2016; Rasouli, Eshuis, Trienekens, Kusters, & Grefen, 2016), several researchers have proposed different data governance models (c.f. Khatri & Brown, 2010; Otto, 2011; Panian, 2010; Weber et al., 2009; Wende, 2007). These researchers help in our understanding of the data governance areas and in shaping its boundaries. However, only a limited number of papers have examined critical success factors (CSFs) for data governance. This research aims to contribute to the body of knowledge by inductively identifying the CSFs for data governance, following the approach of building theory from case study research proposed by Eisenhardt (1989).

The remainder of this research paper is organised as follows. Section 2.4.4 highlights the literature related to data governance, to aid understanding of its concepts. Section 2.4.5 outlines the research methodology and explains the research approach and data gathering and analysis techniques used. Section 2.4.6 presents

the findings and the CSFs identified for data governance, together with detailed descriptions, followed by a section on the possible interconnectedness of the CSFs. Finally, section 2.4.7 presents the concluding remarks and areas of further research.

2.4.4 Background to data governance

Data governance has received much attention in both academic and practitioner communities. The concept has been developed over the last ten years whereby data are considered as valuable assets and as a strategic function within the organisation's structure and are thus placed under corporate governance (Vayghan et al., 2007; Wende, 2007). Data are increasingly being considered as strategic resources and organisations seek to have a data governance programme so that businesses can generate value from their data assets (Khatri & Brown, 2010).

Data governance focuses on who holds the decision rights related to data assets in an organisation (Khatri & Brown, 2010; Otto, 2011) in order to ensure the quality, consistency, usability, security, privacy, and availability of the data (Cohen, 2006; Panian, 2010).

Rau (2004, p. 35) refers to governance as "*the way the organization goes about ensuring that strategies are set, monitored, and achieved*". Horne (1995) connected governance with the optimal use of assets and outlined how data as an asset drives the importance of the governance of data within an organisation. The concept of data as an asset emerged with a report by the Hawley Committee in 1994, which defined data assets as "*data that is or should be documented and that has value or potential value*" (Oppenheim, Stenson, & Wilson, 2003, p. 159). Therefore, the

main driver of data governance is the consideration of data as an asset in an organisation (Panian, 2010).

In our recent study (Chapter 2, paper 1), we analysed academic and practitioner publications on data governance and proposed a data governance activities model. We identified a set of data governance activities that interconnect three main constructs: 1) action (plus) 2) area of governance (plus) 3) decision domain. These activities can be seen in Figure 2-16.

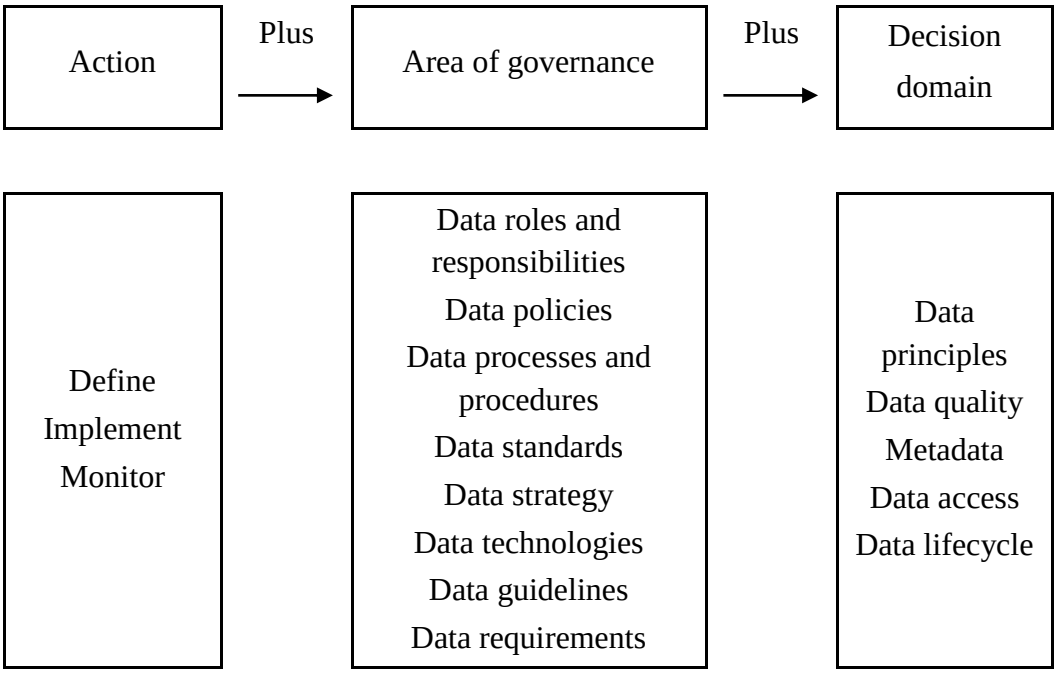

Figure 2-16 illustrates the three data governance activities' constructs, including possible values (Chapter 2, paper 1).

From the above model, it can be seen that there are eight areas of governance reported in the literature across the five decision domains suggested by Khatri and Brown (2010) for which the data governance function should contain decisions around defining, implementing, and monitoring in order to ensure a successful data governance programme. Our analysis shows a lack of research around data

governance, particularly in the implementation and monitoring actions. There is more focus in the literature on the defining action (e.g. define data policies for data quality), which indicates a somewhat embryonic understanding of data governance (Chapter 2, paper 1).

However, while the activities are considered universal data governance activities they can vary from one organisation to another. These can be a guide to understanding data governance-related concepts and boundaries, as they are a collection of the activities reported in different publications. As a result, we needed to understand how governance is actually executed in practice. Therefore, we decided to follow an inductive approach to build theory from a case study. In the next section, we provide a detailed description of our research approach to building theory.

2.4.5 Research methodology

The theory building research strategy proposed by Eisenhardt (1989) provides a clear process for conducting research that aims to build theories from one or more case studies. According to Eisenhardt (1989), the main driver for building theory from a case study is when little is known about a phenomenon and, therefore, the process does not rely on previous literature or prior empirical evidence. Furthermore, interpretive qualitative research is an appropriate research design to apply when exploring CSFs using a case study (Koh et al., 2011). Many scholars have investigated and explored CSFs in certain IS domains and applied qualitative methods using either single case or multiple case study designs (e.g., Butler & Fitzgerald, 1999; Guynes & Vanecek, 1996; Sammon & Adam, 2008).

2.4.5.1 Case background

The Saudi Telecom Company (STC) is the leading telecommunications service provider in Saudi Arabia and the largest telecom provider in the Middle East and North Africa region. It was established on 1998 as the first for-profit telecommunications company in the Kingdom. Before that, telecommunications services were provided by the Ministry of Communication and Information Technology. STC is one of the largest companies listed on the Saudi stock market, with authorised capital of SR 20,000 million ($5,333 million). STC provides all the telecommunications services, such as landlines, mobiles, internet, and television, for the entire country. STC also invests in the telecommunications markets in other countries, such as Indonesia, India, and Kuwait.

We regard STC as an appropriate case for studying a data governance programme for various reasons. Firstly, it is the largest telecom company in the region and serves all the cities and urban areas in Saudi Arabia. Secondly, it provides a full range of telecommunications services, which require different data infrastructures as well as mindsets to manage. Thirdly, the company recently established different services and bundles targeting each customer by studying the behaviour of the customer; this would not have been possible without a proper data governance programme. Within our observations during a series of interview sessions, we can see that a data governance programme is not fully established across every area of the organisation, although many of the data governance practices are well established.

2.4.5.2 Data gathering

The data gathering was guided by the approach outlined by Rockart (1979), who suggests conducting separate interviews with executives individually in order to identify CSFs. According to Rockart (1979), CSFs are "*areas of activity that should receive constant and careful attention from management*" (p. 85). The CSF approach has been widely investigated and used in information systems (IS) research and practice over the last three decades (Shah et al., 2007; Tan et al., 2009) and remains a valid research method for making sense of a problem by identifying potential factors that influence a community of practice (Caralli et al., 2004; Lam, 2005).

This research employed semi-structured interviews, which enabled the researchers to identify and explore the CSFs for data governance. We developed a data collection procedure based on the CSF approach in Rockart (1979) referred to above (see also Figure 2-17 below). Fifteen individual interviews were conducted at STC with personnel at the managerial level of both business and operation departments (see Table 2-20 for a list of the interviewees' positions and the duration of the interviews). These interviews were conducted in two different periods. The researchers decided to stop interviewing more people at the point at which information started to be repeated and the material collected was sufficiently rich to cover the majority of the data governance aspects in the organisation.

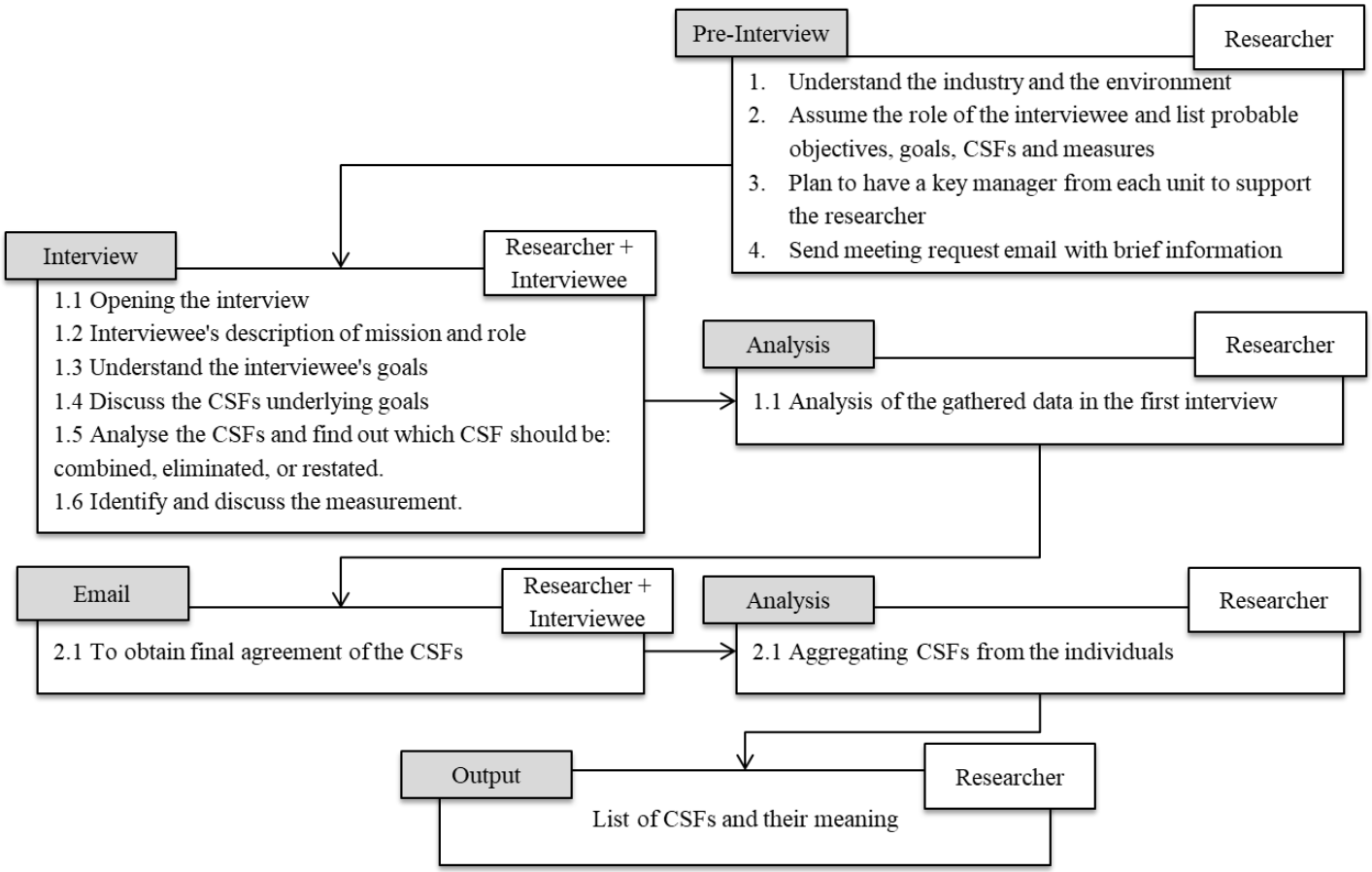

Figure 2-17 Data gathering approach.

All the interviews began with an introduction to the research objective. Each interviewee was then asked to begin talking about the data-related activities in his/her department. Then, during the interviews, we identified the CSFs related to data governance. In many cases, the interviewer explained the data governance programme from the perspective of the five decision domains (c.f. Khatri & Brown, 2010) to ensure the interviewee shared the meaning of data governance. During the interviews, the interviewer attempted to keep the discussion to data-governance-related topics in order to concentrate the interviews around the research subject. Some of the interviews were conducted in Arabic and others in English, depending on the English-language level of the interviewee.

Table 2-20 List of interviewees' positions and related sector and interview duration.

Position	Sector	Interview duration (minutes)
Core and Messaging Services Solutions Director	Operation	60
Voice and Add-ons Director	Business	60
(1) Network Support Application Manager	Operation	45
(2) Network Support Application Manager	Operation	50
Fault Management Systems Manager	Operation	60
Business Partner -HR	Business	60
Unified Communication Section Manager	Operation	80
Messaging and Roaming Services Manager	Operation	45
Sales and Support Director	Business	60
Sales Supervisor - Senior Expert	Business	40
Indirect Sales Manager	Business	40
Virtual Sales Director	Business	60
Sales Support Manager	Business	40
Broadband Services Development Manager	Operation	80
Data Mining and Development Manager	Operation	60
Total 15 interviewees		840 minutes

Table 2-21 illustrates the data preparation steps undertaken for the data to be analysed. All the interviews were transcribed word-for-word and those conducted in Arabic were translated into English by a third party in order to avoid researcher bias. The transcripts were then reviewed with the recording in order to supply any missing words. Finally, the transcripts that had been translated were reviewed to ensure that they were true to the meaning of the original interview.

Table 2-21 Data preparation steps for analysis.

Steps	Role	Description	Status
Record the interview	Researcher	All the interviews were recorded with the prior permission of the interviewees.	Voice only
Transcribe the interview	Third party	The interviews were listened to and the content transcribed in English. There were clear instructions written by the researcher to standardise the method for transcribing the interviews. This was done by a third party in order to avoid researcher bias.	Transcribed with 70% accuracy
Review	Researcher	The researcher listened to the interview recordings again and reviewed the transcript word-by-word to add any missing vocabulary, as well as changing or correcting phrases in order to reflect the meaning when comparing Arabic and English.	Transcribed with 100% accuracy
Data cleaning	Third party	The third party re-read the transcripts, proofread them, and reorganised them into paragraphs in order to make more sense of the data.	Ready for coding

2.4.5.3 Data analysis

Open, axial, and selective coding techniques described by Strauss and Corbin (1990) were adopted in this research in a way that serves the purpose of the research objective. Coding is one of the techniques widely used in analysing qualitative data in order to build theory from a case study (c.f. Buchwald et al., 2014; Tallon et al., 2013; Tan et al., 2015). In the following data analysis, there are, as outlined by Strauss and Corbin (1990), three types of coding: open, axial, and selective (see Table 2-22). These coding techniques are aimed at generating concepts from field data (Walsham, 2006). According to Strauss and Corbin (1990, p. 57), coding "*represents the operations by which data are broken down, conceptualized, and put back together in new ways*".

Table 2-22 Open, axial, and selective coding definitions by Strauss and Corbin (1990).

Coding technique	Definition
Open coding	"*The process of breaking down, examining, comparing, conceptualizing, and categorizing data*" (p. 61).
Axial coding	"*A set of procedures whereby data are put back together in new ways after open coding, by making connections between categories. This is done by utilizing a coding paradigm involving conditions, context, action/interactional strategies and consequence*" (p. 96).
Selective coding	"*The process of selecting the core category, systematically relating it to other categories, validating those relationships, and filling in categories that need further refinement and development*" (p. 116).

Open coding is a process aimed at identifying the concepts or key ideas that are hidden within data that are likely to be related to the phenomenon of interest (Bhattacherjee, 2012). Concepts and categories are generated in the open coding stage (Glaser, 1992). Strauss and Corbin (1990) state that the concepts that appear to be similar are grouped together under a higher-order, more abstract concept called a category.

The second reading of the data is considered during axial coding (Dezdar & Sulaiman, 2009), which is performed simultaneously with open coding (Bhattacherjee, 2012; Strauss & Corbin, 1990). During this stage, the categories are refined in order to be linked in the form of relationships. Strauss and Corbin (1990) suggest that, in order to identify the relationships between data, a paradigm model should be used that consists of causal conditions, the phenomenon, the context, intervening conditions, action/interaction strategies, and consequences. Using this model enables the researcher to think systematically about the data in order to relate them (Strauss & Corbin, 1990).

Finally, selective coding begins when researchers identify a potential core category (Tan et al., 2015), focusing then on the core categories and related categories that emerged during the axial coding. This involves comparing the core categories with the raw data by 'telling the story' of the core categories that emerge (Strauss & Corbin, 1990).

2.4.5.4 Data analysis coding procedure

After preparing all the interview transcripts, the data analysis was commenced by reading each transcript sentence-by-sentence and following an open coding technique. After coding the first two interviews, axial coding was commenced in an iterative manner as categories started to emerge (see Figure 2-18).

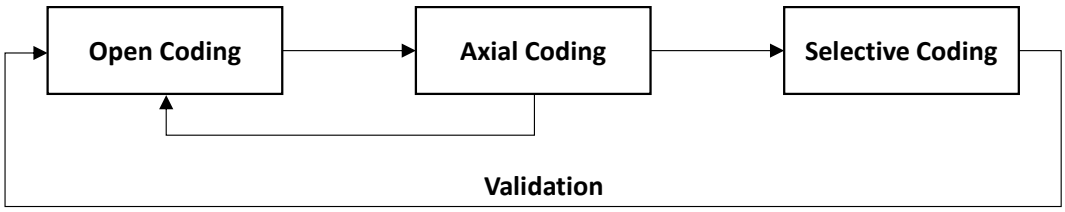

Figure 2-18 Open, axial, and selective coding iterative process.

The five decision domains identified by Khatri and Brown (2010) (see Table 2-23) were used to break the phenomenon down into paradigm models (see Figure 2-19: Paradigm model constructs) in order to clarify the relationships between the categories that emerged during the open coding analysis. The axial coding procedure resulted in five paradigm models that identify the relationships between the categories. The researchers were then able to employ selective coding for the core categories and validate the concepts that emerged in an iterative manner. The core categories are considered later as CSFs for data governance.

Table 2-23 Decision domains for data governance (Khatri & Brown, 2010).

<table>
<tr><td colspan="3">Data principles “Clarifying the role of data as an asset”</td></tr>
<tr><td rowspan="2">Data quality
“Establishing the requirements of intended use of data”</td><td>Metadata
“Establishing the semantics or “content” of data so that it is interpretable by the users”</td><td rowspan="2">Data life cycle
“Determining the definition, production, retention and retirement of data”</td></tr>
<tr><td>Data access
“Specifying access requirements of data”</td></tr>
</table>

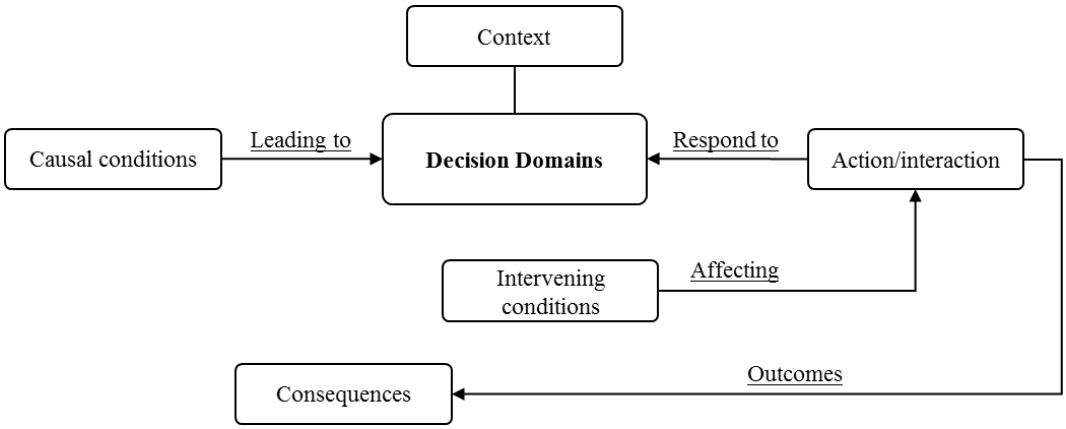

Figure 2-19 Paradigm model constructs.

The coding procedure for the 15 interviews resulted in 325 concepts that related to data governance. The 325 concepts generated 84 categories. Using the paradigm models, the researchers identified the relationships between the 84 categories, which enabled the creation of nine core categories during the selective coding phase. Figure 2-20 illustrates an example the processes, together with examples of the concepts, category relationships, and the core category, namely, ‘Clear, inclusive data requirements’.

Open coding	**Concept #1**: Data owner defines the reporting requirements	**Excerpt**: "*he is looking for a function or a feature, not data; they need the report to be done a certain way, or data in a certain way".*
	Concept #2: Business is not defining the reporting requirements	**Excerpt**: "*the business is not doing their duty the right way. They are not asking for the right information".*
	Concept #3: Be very specific in the requirements	**Excerpt**: "*the more specific you are in your requirements, the more accurate the outcome"*
	Concept #4: Precise data requirements	**Excerpt**: "*The more accurate you are in your requirements, they will implement it correctly".*

Axial coding	**Category #1**: Data reporting requirements	**Concept #1**: Data owner defines the reporting requirements **Concept #2**: Business is not defining the reporting requirements	**Paradigm model** Data reporting requirements are a causal condition for data principles
	Category #2: Clearly specify the data requirements	**Concept #3**: Be very specific in the requirements **Concept #4**: Precise data requirements	**Paradigm model** Specifying data requirements clearly is an action/interaction for data quality

Selective coding	**Core category #1**: Clear, inclusive data requirements	**Cause**: Data reporting requirements
		Actions: Clearly specify the data requirements

Figure 2-20 Examples of the data coding procedure for the 'Clear, inclusive data requirements' CSF.

2.4.6 Findings

From our interviews, and observations made during the interviews, we found that STC did not have a well-established data governance programme for all employees to follow. However, there are many well-defined activities that can be considered as part of a data governance programme, such as having clear data processes and procedures within new systems development, as well as respecting roles and responsibilities in terms of the data. Thus, the business personnel rely heavily on the data provided by the IT department, but there is no opportunity for the business personnel to judge the results.

In the following subsection, a list of the CSFs that emerged from the case study is presented, followed by a description of each. We found nine core categories that can be considered CSFs for data governance from the perspective of telecommunications enterprises.

2.4.6.1 CSFs for data governance

This subsection discusses the CSFs for data governance that emerged from our analysis of the case study. We found nine interrelated core categories that resulted from the selective coding we applied to the paradigm models. These core categories and their properties can be considered CSFs for data governance. The following descriptions of each of the CSFs are ordered based on a frequency count of the concepts they reflect.

CSF 1: Proper data integration strategies

Having proper data integration strategies as a CSF for data governance was referred to by many of the interviewees, so we can argue that this factor plays a critical role in the ability of our case (STC) to increase the success of its data governance programme. This CSF consists of different elements, including a unified database for products and services as well as the overall direction of the data integration. Furthermore, of the elements included, many of the interviewees emphasised the integration between data quality and the advantages presented to the beneficiaries (employees, customers, suppliers, etc.). One of the examples of this is that, in order to have medical insurance, an employee should update his/her information on the human resource (HR) system correctly. As reported by one of the HR managers, "*The employee has to update their information every six months or every year or whenever they have something new, like degrees, CV, kids, or wife. Wife and kids are included in insurance*".

This CSF emerged due to current practices in terms of the integration between services and products within the system infrastructure. This is not only a technical issue, but the strategy and direction of the data integration are a problem within STC. One of the interviewees summarised the current issue by stating: "*We need to do something to unify the database of customers, and then we cleanse it. For example, let's talk about the customers' database. We also have a million databases for other things, but talking about the customers' database, it's too big and too many, with one for cell phone and one for landline, and one for each service*". Hence, the issue becomes critical, with the customer database.

A number of actions are recommended for accomplishing this CSF, which include deciding to make a unique interface that contains all the customer data across the company's products and services. This would enable the company to change practice by assigning an account manager for each customer, instead of the current approach of assigning an account manager for each customer for each service. In addition, data integrity should be integrated with another entity that the person doing the data entry or the manager should consider, thus forcing the beneficiary to insert relevant data to receive a benefit, such as linking the updating of employee data with the insurance offered. Updating and maintaining network data are also connected with employees' key performance indicators (KPIs).

CSF 2: Employee data competencies

The employee data competencies factor covers data governance activities that involve human activity, as well as employees' skills and capabilities. The competencies of all employees, from senior executives to entry-level workers, are important due to their involvement in various data governance activities at various points in time. For example, during a discussion on the topic of dealing with different types of data, one of the interviewees stated: "*We have experts in each domain, and in each domain we have specialists for each system*". At the executive level, establishing an overall data governance strategy requires top managers to have certain competencies. Based on our analysis, such competencies would be needed to treat data as a strategic asset.

Employee data competencies are considered as CSFs due to the direct involvement of the employees during the data life cycle, starting with data entry and including

data governance activities. For example, within the context of the reasons for not having a proper data governance programme, one of the interviewees stated: "*but I see that the issues mainly come from the human use, not from the systems*". Dealing with data entry and data access also requires employees to have a minimum set of capabilities and a certain level of awareness in order to handle the organisation's data. For example, the practice of manual data entry depends on a certain level of employee capabilities and awareness of data processes and procedures.

In order to ensure appropriate employee data competencies, the most important actions are 'training' and 'awareness', such as continuous training in dealing with and implementing data processes and procedures, and includes internal and external training. Many of the interviewees emphasised the need to educate the people who are dealing with the data, starting at the data entry level. For example, as stated by one interviewee, "*As for the solution, I believe the reason for the problem is the people, so we need to educate the people*". It is also vital to increase employees' awareness of the criticality of data in terms of entering the right data, as well as when accessing and sharing sensitive data.

CSF 3: Flexible data tools and technologies

Flexible data tools and technologies were found to be one of the highest reported critical factors for governing data successfully. This includes the different elements that enable some of the other CSFs to be performed successfully, and, for example, the data integration infrastructures that enable the better use of data due to the different services provided to each customer. For example, as stated by one of the interviewees: "*We have fibre, access, mobile, many things. If you work with each*

and every one, this is not handy, as some of them are very old and you can't get benefit from them". Hence, data integration is recommended in order to have accurate information about each customer across all services.

Considering the flexible data tools and technologies as CSFs for data governance is not surprising in the telecommunications industry. Today, technologies relating to telecommunications services are rapidly changing and being updated with more innovative features. Hence, telecommunications companies should keep updating new services in order to compete in the market. However, frequent updating requires flexible tools and technologies, in order to migrate and integrate data within different systems.

Having appropriate IT infrastructure and integrated data are recommended to address the need for flexible data tools and technologies. This includes achieving a strategy by setting up advanced technologies that enable data integration in order to automate the validation of the data. It is also recommended that systems are able to deal with live big data to enable the micro-segmentation of customer behaviour to support the business need for monitoring customers' activities, as stated, for example, by one of the interviewees: "*So always do the micro-segmentation, and we can't do it unless we have a strong data warehouse*". Finally, it is important to take into account the privacy and availability of data while integrating internal and external systems.

CSF 4: Clear, inclusive data requirements

Data requirements play a critical role in any data governance programme and include those relating to gathering behaviour, which consists of the use of

standardised language and methods for collecting data. This process shapes all the business requirements within the data flow and presentation. Our analysis shows that many of the interviewees focus on data presentation requirements, including analytics. For example, one interviewee reported on the context of data reporting as follows: "*Sometimes it does not make sense, it gives a wrong reading, but the data is correct and it's just the way of analysing it*".

Our analysis shows that one of the elements of the data requirements highlighted are the communication processes within a project. These processes clarify the data requirements so that they can be understood by the implementers and the correct implementation of the requirements is followed. As stated by one of the data business owners, "*The more accurate you are in your requirements, they will implement it correctly*". Therefore, from our analysis, we find that there is a clear two-way communication procedure embedded into the project management tools within the organisation. This enables the requesters (business data owners) to ensure that their requirements are communicated well and implemented in the right way.

A number of actions are recommended within the 'Clear, inclusive data requirements' CSF that are mainly associated with employee knowledge and behaviour when establishing new requirements by stating detailed and comprehensive conditions related to data storage, structure, and presentation. However, in contrast, failing to have clear data requirements opens the door for the implementers to substitute any missing instructions in accordance with their scope of understanding. As stated by one of the interviewees in the IT sector, they work

"*According to what we get, and sometimes we got unclear requirements, so if it wasn't explained to us we just implemented it the way we understood it*".

CSF 5: Clear data processes and procedures

Clear data processes and procedures are considered one of the CSFs for data governance among telecommunications companies. This seems obvious, as the nature of telecommunications companies means they deal with a very high volume of data coming from different sources. In order to govern these data, a company should have clear processes and procedures within the data life cycle stages.

Having clear data processes and procedures enables the building of trust in data quality, as the beneficiaries of the data know exactly how the data are processed. For example, one of the interviewees clearly stated that "*Whenever you have processes to maintain, you will be fine with the data*".

Embedded data processes and procedures are considered to be the main recommended action for this CSF, due to the high volume of data entries and the thousands of employees entering the data initially. Hence, these processes and procedures should be embedded in the systems to eliminate the possibility of entering inaccurate data. For example, as stated by one of the interviewees, "*a certain process is defined, but it has to be translated into applications, not just on paper*".

CSF 6: Focused and tangible data strategies

Focused and tangible data strategies include planning for data governance in order to achieve its goals, as well as the main activities related to considering data as

assets. Furthermore, this includes the short- and long-term objectives that relate to data governance. Furthermore, our analysis shows that the return on investment in data should be calculated in order to persuade the top managers to consider data as assets.

One of the main causal conditions for requiring this CSF is the practice of not considering data as a strategic element, as this indicates that executives do not recognise the value of such governance. For example, one interviewee stated that "*Data governance and data integrity are tools to help reach the right approach, in their plan for higher revenues. Their revenues are going up and down and affecting the marketing, so they are focusing on this and their least priority is this governance part*".

Some actions are recommended in order to address this CSF include understanding the importance of data and considering them as assets. For example, as stated directly by one of the interviewees, "*Data is the most important thing in any organisation and any system*". In addition to the recommended actions is calculating the return on investment and considering the consequences of having a proper data governance programme.

CSF 7: Established data roles and responsibilities

Established data roles and responsibilities should also be considered in the context of data governance. It is important to identify the individual(s) responsible for the data-related activities in an organisation, such as who should be accountable for the data processes and the requirements for the data, as well as assigning duties in relation to undertaking actions related to data. From a strategic point of view,

establishing data roles and responsibilities is included in the data governance function in an organisation, as well as defining who owns the data.

Our analysis shows that data roles and responsibilities are established within STC and we see that many of the interviewees indicate this by mentioning some of the roles and responsibilities involved. For example, one of the interviewees who benefits from data stated: "*From my side, I'm not in the position to judge their way. They give me the data and they are responsible for that*". However, although there are clear data roles and responsibilities, this does not mean the data are trusted in terms of quality. Many of the interviewees argued that the people responsible for the data are not doing their job as they should, due to their responsibilities not being outlined fully in their job description and associated with their KPIs, particularly in the context of the level of the data quality. For example, in the context of data quality, one interviewee stated that "*The current problem is that when the job description is not well established, he doesn't know what he should do. Currently it has only a general statement, which says that you need to do anything related to your work*".

From our analysis, we recommend setting up a committee for data governance, as well as identifying data owners. For example, as stated by one interviewee, "*the suggestion for data governance success is that there should be a data governance owner, data owners or a high committee between the business and the IT and people who are involved in it should be from different backgrounds, IT and business*". It is recommended, therefore, to assign a committee to deal with all data governance activities.

CSF 8: Accountable data access and availability

Accountable data access and availability consists of data access privileges, as well as data availability at the right time and in the right format. Our analysis shows emphasis on this CSF from different interviewees, as a telecommunications company deals with sensitive and personal data. Hence, having a policy for data access is required to prevent data leakage.

One of the conditions that causes this CSF to accrue and be considered part of a successful data governance programme is the availability of data for the decision maker at the right time and in the right format, which enables better decisions. For example, when targeting the right customers for promotions relating to current services based on their service usage, one of the marketing managers reported the following: "*When you design promotions or services, they know the targeted segment, and know if they are heavy users of the internet, so the designer of the services does segmentation*". Therefore, having data available in the right format enables marketing activities segmentation.

Data availability should also be aligned with the data access regulations to prevent data leakages. Therefore, as personnel are required to be accountable for and safeguard data assets. Additional recommended actions are defining who can access the data as well as implementing the right technologies to enable data to be available at the right time and in the correct format.

CSF 9: Effective data monitoring and feedback

Effective data monitoring and feedback consists of data auditing and tracking. In order to conduct a data governance programme successfully, data should be monitored and audited. Our analysis stresses the need to implement a data auditing tool. For example, one of the interviewees within the context of governing data reported that "*each system in the network has a performance tool connected to it*".

This CSF is accrued as the feedback from the data monitoring team is used for the continuous improvement of data quality. This is important for telecommunications enterprises, as they provide and sell data in the form of calls, data, and text messages. These services require continuous monitoring and improvement of the accuracy of the data concerned, an example of which was provided by one of the interviewees, who described "*monitoring data, detecting team, a big team with authority to monitor data from A to Z and revisit data from time to time in order to make it accurate*".

However, having effective data monitoring and feedback requires certain tools and technologies that enable the data to be tracked and auto-auditing activities and improvement to be supported, as well as monitoring systems that possess proper data.

2.4.6.2 Possible interconnectedness of CSFs

As an outcome of our analysis, the possible interconnectedness of the CSFs is also explored. This enables a better understanding of the multiplicity of effects of each CSF. The interconnections articulate the existing impact of the absence/presence of

one of the CSFs on each other. Table 2-25 presents the possible interconnectedness between the CSFs. However, the absence of one of the CSFs not having a direct impact on other CSFs, such as 'Proper data integration strategies', does not mean that the CSF is of lower importance compared with the others. The potential interconnectedness only shows the possible impact of the presence or absence of one CSF on another.

Table 2-24 Possible interconnectedness of CSFs.

CSF	Has an impact on	Interconnection
Employee data competencies	Established data roles and responsibilities	Defines the competency requirements for each role in order to perform the other CSFs
Flexible data tools and technologies	Proper data integration strategies	Enables the implementation of data integration strategies
	Clear data processes and procedures	Embeds the data processes and procedures into the systems
	Effective data monitoring and reporting	Enables the establishment of data monitoring and reporting tools
Clear, inclusive data requirements	Flexible data tools and technologies	Defines the data tools and technologies required
Established data roles and responsibilities	All CSFs	Defines who does what and assigns responsibilities for each activity

From Table 2-25, it can be seen that the CSF ‘Employee data competencies’ has a direct impact on ‘Established data roles and responsibilities’, as each data role and responsibility requires certain employee competencies in order to be performed successfully. In turn ‘Established data roles and responsibilities’, has a direct impact on the other CSFs, as the absence of data roles and responsibilities has a direct impact on conducting the other CSFs. Therefore, without paying attention to the level of importance of the CSFs reported within our analysis, the actions related to the CSF ‘Established data roles and responsibilities’ should be considered the first action to be taken in order to perform the other CSFs.

In addition, ‘Flexible data tools and technologies’ has an impact on three of the CSFs, as shown in Table 2-25, due to the governing of data relying heavily on the technologies that are involved. For example, ‘Proper data integration strategies’ is impacted along with the current tools and technologies, as these strategies need to be implemented within the IT infrastructure. Finally, ‘Flexible data tools and technologies’ is interconnected with ‘Clear, inclusive data requirements’, as this defines the data tools and technologies required within the data governance programme.

2.4.7 Concluding remarks and future research

This research has attempted to contribute to the knowledge base by identifying CSFs for data governance. This research followed a theory building technique by conducting several semi-structured interviews guided by the CSF approach within a single case study organisation. The interviews were transcribed and prepared for analysis by applying open, axial, and selective coding.

Nine CSFs emerged from the single case study: 1) Proper data integration strategies; 2) Employee data competencies; 3) Flexible data tools and technologies; 4) Clear, inclusive data requirements; 5) Clear data processes and procedures; 6) Focused and tangible data strategies; 7) Established data roles and responsibilities; 8) Accountable data access and availability; and 9) Effective data monitoring and feedback.

The CSFs identified above were ordered depending on the frequency count (number of associated concepts) reported by the interviewees. We found that 'Proper data integration strategies' was the highest reported CSF and was referred to by all the interviewees, albeit with different emphases, due to the current issue in the company related to data integration, particularly customers' data. Our analysis shows that the interviewees associate the success of a data governance programme within their organisation with decisions related to data integration initiatives that lead to a proper data governance programme. We also found that 'Employee data competencies' was considered to be one of the highest-ranked critical factors for successful data governance. This is not surprising, as employees handle all the other factors and activities related to data governance. Therefore, considering employees' competencies, including their skills, training, and awareness, is vital for the success of the other CSFs. In addition, we highlighted the possible interconnectedness of CSFs, which articulates the existence of an impact of the absence/presence of one of the CSFs on each of the others.

Certain limitations are present in any piece of research. The main limitation of this research is that the CSFs reported came from a single case study, which means that

they are may not represent an extensive list of CSFs for data governance. The CSFs identified may, however, be useful in providing guidelines for those who want to conduct a data governance programme within a large telecommunications enterprise. Hence, it is recommended that further research examines the CSFs and the activities involved within each one in order to better understand the implementation road map of each factor.

2.4.8 References

Bhattacherjee, A. (2012). Social science research: principles, methods, and practices (2nd ed.). Tampa, FL: CreateSpace Independent Publishing Platform.

Buchwald, A., Urbach, N., & Ahlemann, F. (2014). Business value through controlled IT: toward an integrated model of IT governance success and its impact. Journal of Information Technology, 29(2), 128-147. Retrieved from https://link.springer.com/article/10.1057/jit.2014.3

Butler, T., & Fitzgerald, B. (1999). Unpacking the systems development process: an empirical application of the CSF concept in a research context. The Journal of Strategic Information Systems, 8(4), 351-371. doi:10.1016/S0963-8687(00)00027-5

Caralli, R.A., Stevens, J.F., Willke, B.J., & Wilson, W.R. (2004). The critical success factor method: establishing a foundation for enterprise security management. Retrieved from http://www.dtic.mil/get-tr-doc/pdf?AD=ADA443742

Cohen, R. (2006). What's in a name? Data governance roles, responsibilities and results factors. DM Review, 8.

Dezdar, S., & Sulaiman, A. (2009). Successful enterprise resource planning implementation: taxonomy of critical factors. Industrial Management & Data Systems, 109(8), 1037-1052. doi:10.1108/0263557091099

Eisenhardt, K.M. (1989). Building Theories from Case Study Research. The Academy of Management Review, 14(4), 532-550. doi:10.5465/AMR.1989.4308385

Glaser, B. G. (1992). Emergence vs forcing: Basics of Grounded Theory Analysis. Mill Valley, CA: Sociology Press.

Guynes, C.S., & Vanecek, M.T. (1996). Critical success factors in data management. Information & Management, 30(4), 201-209. doi:10.1016/0378-7206(95)00053-4

Hassan, S., & Chindamo, P. (2017). Governance in practice: Effective data governance: From strategy through to implementation. Governance Directions, 69(4), 207.

Horne, N.W. (1995). Information as an asset—The board agenda. Computer Audit Update, 1995(9), 5-11. doi:10.1016/0960-2593(95)90246-5

Kamioka, T., Luo, X., & Tapanainen, T. (2016). An empirical investigation of data governance: The role of accountabilities. The 20th Pacific Asia Conference on Information Systems (PACIS),

Khatri, V., & Brown, C.V. (2010). Designing data governance. Communications of the ACM, 53(1), 148-152. doi:10.1145/1629175.1629210

Koh, S.C.L., Gunasekaran, A., & Goodman, T. (2011). Drivers, barriers and critical success factors for ERPII implementation in supply chains: A critical analysis. The Journal of Strategic Information Systems, 20(4), 385-402. doi: 10.1016/j.jsis.2011.07.001

Lam, W. (2005). Investigating success factors in enterprise application integration: a case-driven analysis. European Journal of Information Systems, 14(2), 175-187. Retrieved from https://link.springer.com/article/10.1057/palgrave.ejis.3000530

Mohammad, R. R., Jos, J. M. T., Kusters, R. J., & Paul W.P.J. Grefen. (2016). Information governance requirements in dynamic business networking. Industr Mngmnt & Data Systems, 116(7), 1356-1379. doi:10.1108/IMDS-06-2015-0260

Nagle, T and Sammon, D, (2017). The Data Value Map: A Framework For Developing Shared Understanding On Data Initiatives. In Proceedings of the 25th European Conference on Information Systems (ECIS), Guimarães, Portugal, June 5-10, 2017 (pp. 1439-1452). ISBN 978-989-20-7655-3 Research Papers. https://aisel.aisnet.org/ecis2017_rp/93

Oppenheim, C., Stenson, J., & Wilson, R.M.S. (2003). Studies on Information as an Asset I: Definitions. Journal of Information Science, 29(3), 159-166. doi:10.1177/01655515030293003

Otto, B. (2011). Organizing data governance: findings from the telecommunications industry and consequences for large service providers. Retrieved from http://aisel.aisnet.org/cgi/viewcontent.cgi?article=3610&context=cais

Panian, Z. (2010). Some practical experiences in data governance. World Academy of Science, Engineering and Technology Management, 62, 468-475.

Rau, K. G. (2004). Effective governance of it: Design objectives, roles, and relationships. Information Systems Management, 21(4), 35-42. doi:10.1201/1078/44705.21.4.20040901/84185.4

Rockart, J.F. (1979). Chief executives define their own data needs. Harvard business review, 57(2), 81-93. Retrieved from http://europepmc.org/abstract/med/10297607.

Sammon, D., & Adam, F. (2008). Justifying an ERP investment: critical success factors for transformation investments. ICIS 2008 Proceedings, 60.

Shah, M.H., Braganza, A., & Morabito, V. (2007). A survey of critical success factors in e-Banking: an organisational perspective. European Journal of Information Systems, 16(4), 511-524. Retrieved from https://link.springer.com/article/10.1057/palgrave.ejis.3000693

Strauss, A., & Corbin, J. (1990). Basics of qualitative research: Grounded theory procedures and techniques. Thousand Oaks, CA: SAGE Publications, Inc.

Tallon, P.P., Ramirez, R.V., & Short, J.E. (2013). The Information Artifact in IT Governance: Toward a Theory of Information Governance. Journal of Management Information Systems, 30(3), 141-178. doi:10.2753/MIS0742-1222300306

Tan, B., Pan, S.L., Lu, X., & Huang, L. (2015). The Role of IS Capabilities in the Development of Multi-Sided Platforms: The Digital Ecosystem Strategy of Alibaba. com. Journal of the Association for Information Systems, 16(4), 248-280.

Tan, W.G., Catersteel, A., & Toleman, M. (2009). Implementing it service management: A case study focussing on critical success factors. Journal of Computer Information Systems, 50(2),1-12. doi:10.1080/08874417.2009.11645379

Vayghan, J. A., Garfinkle, S. M., Walenta, C., Healy, D. C., & Valentin, Z. (2007). The internal information transformation of IBM. IBM Systems Journal, 46(4), 669-683. doi:10.1147/sj.464.0669

Walsham, G. (2006). Doing interpretive research. European journal of information systems, 15(3), 320-330. Retrieved from https://link.springer.com/article/10.1057/palgrave.ejis.3000589

Weber, K., Otto, B., & Österle, H. (2009). One Size Does Not Fit All--A Contingency Approach to Data Governance. Journal of Data and Information Quality (JDIQ), 1(1). doi:10.1145/1515693.1515696

Wende, K. (2007). Proceedings from 18th Australasian Conference on Information Systems (ACIS): A model for data governance–Organising accountabilities for data quality management. Toowoomba, Australia: The University of Southern Queensland.

Chapter Three

3. Discussion and Conclusion

3.1 Introduction

Building on the previous chapters, this chapter combines the results from the two case studies to formulate a final list of CSFs for data governance. The final CSFs are explained in detail and associated with actions recommended from our analysis. The research objective is *"to identify the critical success factors for data governance that enable organisations to introduce an effective data governance programme"*. This is achieved by answering the following research questions:

RQ1: *What are the CSFs for data governance?*

RQ2: *What are the recommended actions for putting the CSFs for data governance into practice?*

This chapter also considers why these factors are critical for data governance and the possible interconnectedness of the CSFs, which articulates the impact of the absence/presence of one of the CSFs on the others.

The remainder of this chapter is organised as follows. Section 3.2 presents a comparison between the CSFs for data governance identified in each case study and includes a final list of the CSFs for data governance that resulted from the both case studies. Section 3.3 answers RQ1 by providing a full description of each of the nine CSFs identified. This is followed by section 3.4, which aims to answer RQ2 and

includes the recommended actions for putting the nine CSFs into practice. Section 3.5 explores the possible interconnectedness of the various CSFs. Section 3.6 discusses and compares the actions associated within a data governance programme reported in the literature with the CSFs for data governance identified in this research study. Section 3.7 presents concluding remarks for the entire research study and a summary of the findings. The subsections include the theoretical contributions of the research, its practical contributions, limitations and suggestions for future research.

3.2 Comparison between the Case Studies

In this section, a comparison between the CSFs reported in each case study outlined in the previous chapter (Chapter 2, papers 3 and 4) is presented. The first case (Alrajhi Bank) identified seven CSFs for data governance, and the second (STC) identified nine. When compared, six of the CSFs were identified in both cases: 1) Employee data competencies, 2) Clear data processes and procedures, 3) Flexible data tools and technologies, 4) Established data roles and responsibilities, 5) Clear, inclusive data requirements, and 6) Focused and tangible data strategies. The remaining CSFs were identified in one of the cases. These CSFs are as follows: from the first case, 7) Standardised easy-to-follow data policies; from the second case, 8) Proper data integration strategies, 9) Accountable data access and availability, and 10) Effective data monitoring and feedback.

The difference between the CSFs identified in each case is due to the data maturity level and the motivation for having an effective data governance programme within each case. For example, in the second case (STC), the top CSF is ‘Proper data

integration strategies', which was not reported by the first case (Alrajhi Bank). As explained previously (see Paper 4), there is a clear problem within STC with regard to its integration practices of the business data and it requires a strategic direction to solve this issue, thereby needing an effective data governance programme. Therefore, this was considered the highest CSF prioritised in STC.

However, the 'Accountable data access and availability' CSF reported by the second case (STC) can be considered part of the 'Standardised easy-to-follow data policies' CSF reported in the first case (Alrajhi Bank). This is because there is a focus on data policies by Arajhi Bank due to the nature of banking data, which includes data access and availability policies. However, in the second case (STC), there is no focus on the data policies in general. The focus is only on the data access and availability policies and these are not reported very frequently. Therefore, within the cross-case analysis, it was decided to include the 'Accountable data access and availability' CSF within the 'Standardised easy-to-follow data policies'.

The final list of CSFs is illustrated in Figure 3-1 and includes nine CSFs for data governance. The CSFs in the figure are ranked based on the total number of concepts included in each of the CSFs accumulated across both case studies. In addition, Figure 3-1 also shows the percentage of each CSF associated with each case. For example, for CSF 1: Employee data competencies, 44% of the concepts included in this CSF are coded from STC and 56% from Alrajhi Bank.

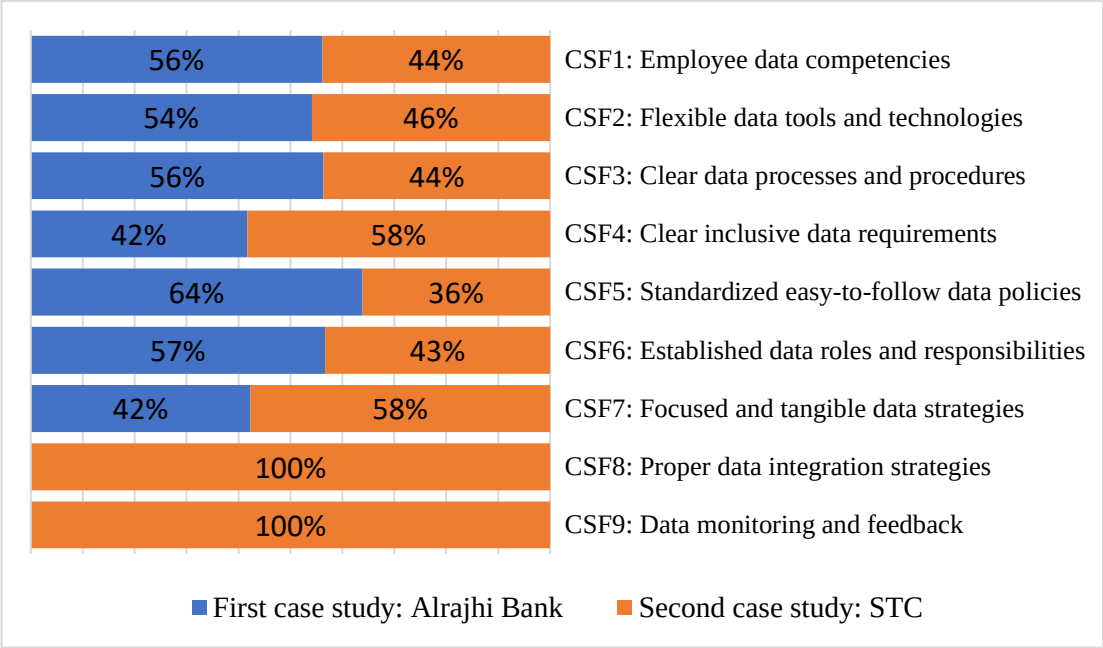

Figure 3-1 The CSFs identified, including the percentage of each CSF in each case study.

In general, of the CSFs identified in both cases, it is evident that, the majority of the CSFs receive more focus in the first case (Alrajhi Bank). However, two CSFs (CSF 8 and CSF 9) are only associated with the second case. Furthermore, CSF 4, 'Clear, inclusive data requirements', has more focus from the second case (STC) in comparison with the first (Alrajhi Bank). From our observations during the two case studies, the data requirements are clearly identified and have matured in the first case (Arajhi Bank) as they are part of their internal regulations.

In addition, there was a lack of focus on the strategic elements related to data governance in the first case (Alrajhi Bank) in comparison with the second (STC). It can be seen that the CSF 'Focused and tangible data strategies' was ranked as the least important CSF for data governance (see Chapter 2, paper 3). Although there

were strong statements by some of the interviewees, the majority did not mention any kind of strategic elements needed for governing data. On the other hand, in the second case (STC), the majority of the interviewees discussed some of the strategic elements related to data governance and, due to considerable focus on the strategic elements, 'Proper data integration strategies' are considered as an independent CSF and ranked as the highest reported CSF.

Ultimately, we can argue that, from both case study perspectives, the CSFs related to employee data competencies and data roles and responsibilities, as well as data tools and technologies, are equally important and all receive the same focus. However, it can also be argued that the first case (Alrajhi Bank) focuses more on the regulations and policies for data governance, whereas the second case (STC) has greater focus on the strategic elements of data governance. However, both perspectives were driven by current practices and the need to govern data.

The following section presents a detailed explanation of the nine CSFs for data governance identified within the cross-case analysis. As part of the explanations, an attempt is made to outline the following: What are the CSFs? Why are they critical? and What are the recommended actions for each CSF?

3.3 CSFs for Data Governance

3.3.1 CSF #1: Employee data competencies

The ‘Employee data competencies’ factor covers data governance activities that involve human activity, as well as employees’ skills and capabilities. This factor includes the competencies of all employees who are involved in the various data governance activities, from senior executives to entry-level workers. Therefore, employee data competencies should be identified for each of the established data roles and responsibilities, in order for the nominated employee to undertake the task successfully.

The ‘Employee data competencies’ factor is considered the most critical from both case studies, due to the involvement of employees in executing all the activities required for data governance, as well as the direct impact this has on the other CSFs, as it determines an employee’s ability to handle the actions specified. For example, establishing an overall data governance strategy requires certain upper-level managers to have certain competencies. Such competencies would also be needed to treat data as a strategic asset. Furthermore, dealing with data entry and access also requires employees to have a minimum set of capabilities and a certain level of awareness with regard to handling the organisation’s data.

In order to ensure appropriate employee data competencies, a number of actions are recommended following the analysis. The recommended actions start with defining the required skills and competencies for dealing with the data for each data role and responsibility. This enables the decision maker to nominate the right person for the

right position, as well as facilitating actions related to implementing the required competencies. The implementation action is followed by conducting training for those employees in relation to the data policies and procedures, as well as increasing awareness of the importance of data and considering them as part of the organisation's assets. Finally, in terms of monitoring these actions, it is recommended to monitor employee activities and performance regarding data in order to plan the right training and encourage awareness. Table 3-1 summarises the actions recommended for the three actions ('define', 'implement', and 'monitor') required for increasing employee data competencies.

Table 3-1 Summary of the actions recommended for the 'Employee data competencies' CSF.

Recommended actions for 'Employee data competencies'	
Define	The required skills and competencies for dealing with the data for each role and responsibility.
Implement	Education and training for employees in 'how to deal with data', as well as 'increasing awareness of the importance of data'.
Monitor	Employee activities and performance when using data.

3.3.2 CSF #2: Flexible data tools and technologies

Flexible data tools and technologies are intended to handle all the technical elements within a data governance programme, including software and hardware, that enable data processing, storage, integration, and presentation. This CSF also includes the different elements that empower other CSFs to be successful through embedding them into systems.

‘Flexible data tools and technologies’ is considered to be one of the highest-reported CSFs for having a proper data governance programme, as it directly impacts the performance of other activities within each CSF, such as enabling the embedding of data processes and procedures into the business systems, as well as embedding data policies that include monitoring user activities. From our analysis, this CSF occurred due to issues that relate to data integration and legacy systems. In addition, flexible data tools and technologies are required in order to address data entry errors by enforcing the correct data format and values.

Hence, many of the actions that are recommended from our analysis address CSF #2 (see Table 3-2). These actions include defining the data life cycle requirements, together with the data format and metadata, to ensure the consistency and readability of the data, as well as meeting the technical data integration needs. Embedding data processes and procedures is also recommended to eliminate the possibility of entering incorrect data, as well as implementing an appropriate technical architecture for meeting data integration requirements. Finally, it is recommended to have data tools that monitor data use, data life cycle, and data leakage to minimise the chances of data losing value.

Table 3-2 Summary of the actions recommended for the ‘Flexible data tools and technologies’ CSF.

Actions recommended for ‘Flexible data tools and technologies’	
Define	Data life cycle requirements to do for integration technical needs.
Implement	An appropriate technical architecture to meet integration and the data life cycle needs.
Monitor	Data life cycle, and data integration.

3.3.3 CSF #3: Clear data processes and procedures

‘Clear data processes and procedures’ is considered as a CSF for data governance, particularly in large organisations that deal with a high volume of data. Data processes and procedures include all the detailed activities related to data capture, retrieval, flow, integration, authorisation processes, validation, and more, which are related to the data life cycle through the system and manual procedures.

The employment of clear data processes and procedures enables the building of trust in data quality, as the beneficiaries of the data know exactly how the data are processed. However, the absence of data processes and procedures, resulting from a lack of data integrity, can cause part of the data processes and procedures to be missed, such as data testing. Our analysis also shows that one of the reasons for having low data quality is a lack of data processes and procedures for users who entering data manually, as well as the validation of the data in the system.

Hence, in order to address this CSF, some actions are recommended while it is employed (see Table 3-3). The actions include defining how the data are captured, either manually or from another system, as well as specifying the data retrieval processes. In addition, embedding data processes and procedures into the business systems is recommended in order to automate data validation and force data entry users to enter the right data in the correct format. Furthermore, data processes and procedures should be defined by the data owner of each data set in order for this definition to be implemented and monitored correctly.

Table 3-3 Summary of the actions recommended for the 'Clear data processes and procedures' CSF.

Actions recommended for 'Clear data processes and procedures'	
Define	Data capture and retrieval processes for all organisational data.
Implement	Data capture and automated validation by embedding them into business systems.
Monitor	The data flow and data use.

3.3.4 CSF #4: Clear, inclusive data requirements

Data requirements play a critical role in any data governance programme; these are the requests that are initially made by business to IT with regard to data. Data requirements include those relating to the gathering method, which consists of the use of standardised language and templates for building business data requirement documents. This process shapes all the business requirements within the data flow and presentation. It also includes the communication practices between the data owners and the systems' developers.

The 'Clear, inclusive data requirements' CSF accrued due to the need to clarify the communication between the data business owner and the systems' developers with related to data needs. The data requirements also impact the implementation actions of the other CSFs, as these actions need to be defined in standards and structured template, as well as communicated properly with the developers in order to address the other CSFs successfully.

Different actions are recommended to address this CSF (see Table 3-4). These actions include those associated with employee knowledge and behaviour when establishing new requirements, by stating detailed and comprehensive conditions related to data storage, structure, and presentation. In many cases, the defining of data requirements should also involve data access and availability policies, as well as the data life cycle. These elements can be ensured by defining data requirement standards and structure templates. Data requirements also need to be tracked and monitored by continuous communication between all the parties concerned, in order to ensure the correct implementation of the data.

Table 3-4 Summary of the actions recommended for the 'Clear, inclusive data requirements' CSF.

Actions recommended for 'Clear, inclusive data requirements'	
Define	Data requirement standards and structure template.
Implement	An appropriate data requirement standards template.
Monitor	The communication between parties regarding data requirements specification.

3.3.5 CSF #5: Standardised easy-to-follow data policies

'Standardised easy-to-follow data policies' play a fundamental role as a CSF for data governance. Data policies are short statements that define the high-level guidelines and rules necessary for dealing with data. In the context of data governance, data policies include the policies for data regulations, as well as data

access rights and privacy. In addition, the internal and external regulations for data should be addressed within the scope of the data policies.

Standardised easy-to-follow data policies have accrued as a CSF for data governance as, within the data governance context, data are considered assets that are required to be treated carefully and safeguarded. This can be achieved by considering standardised data policies and by making sure there is no leakage of data by allowing access to the data without following well-established data policies.

Our analysis suggests that different actions are recommended to put this CSF into practice (see Table 3-5). The actions include having a strong, clear, simple, and easy-to-follow data policy that follows standards that fit the organisation's needs, and a definition of the data regulations, access rights and data privileges. Considering the implementation of the data policies by embedding them into the business system is also recommended, such as by creating different data access policies with different privileges to ensure data are safe and available when needed. Finally, internal and external data regulation within data policies are recommended to enable the monitoring of data and to ensure compliance with the regulators.

Table 3-5 Summary of the actions recommended for the 'Standardised easy-to-follow data policies' CSF.

Actions recommended for 'Standardised easy-to-follow data policies'	
Define	The data regulations, access rights and privileges.
Implement	The data policies within the business systems.
Monitor	The compliance with external and internal data regulations.

3.3.6 CSF #6: Established data roles and responsibilities

‘Established data roles and responsibilities’ mainly include identifying the individual(s) responsible for the data-related activities in the organisation, such as who defines the policies and processes for the data, as well as assigning duties regarding the actions related to the data. In addition, from a strategic point of view, established data roles and responsibilities include the data governance function position in the organisation, as well as defining who owns the data.

Many causal conditions confirm ‘Established data roles and responsibilities’ as a CSF for data governance. For example, having good processes in place without clear roles and responsibilities leads to mistakes in dealing with data. Imprecise roles and responsibilities caused either by the roles being unclear or having unclear assignments can also result in an overlapping between the tasks related to data.

Actions recommended to address the above CSF include defining the data governance position and data decision rights by identifying the roles and responsibilities for data-related tasks, as well as setting the boundaries and scope of the data owners’ responsibilities. It is also recommended that the assignment of data-related roles to the appropriate decision areas is implemented, as well as ensuring the clarity of the assignments for each of the roles. Table 3-6 summarises the actions recommended for addressing this CSF.

Table 3-6 Summary of the actions recommended for the 'Established data roles and responsibilities' CSF.

Recommended actions for 'Established data roles and responsibilities'	
Define	The data governance position and data decision rights.
Implement	The assignment of data roles to decision areas.
Monitor	The clarity of data responsibilities.

3.3.7 CSF #7: Focused and tangible data strategies

The 'Focused and tangible data strategies' CSF includes planning for data governance in order to achieve its goals, as well as ensuring that the main activities related to considering data as assets, including the short- and long-term objectives that relate to data governance, are included. Consideration of the return on investment in data assets should be part of the data strategy.

Understanding the importance of data and considering them as assets confirms 'Focused and tangible data strategies' as a CSF for data governance. Hence, the governance of the data should drive value for the organisation and a clear return on the investment in its data assets. Our analysis also shows that poor planning negatively impacts data when focused and tangible data strategies are absent.

Considering data as strategic elements is the main action recommended for addressing focused and tangible data strategies. This can be achieved by defining the data value and objectives. Top management enforcement should also be taken as an action and includes considering the assignment of a top committee for data

governance. In addition, it is recommended to review the return on investment in data assets, as well as the value of the data for the organisation, in order to enable better decisions related to these assets. Finally, an overall data governance model should be implemented to enable strategic data plans to be converted into tactical plans. Table 3-7 summarises the actions recommended for the define, implement, and monitor actions.

Table 3-7 Summary of the actions recommended for the 'Focused and tangible data strategies' CSF.

Actions recommended for 'Focused and tangible data strategies'	
Define	Data value and objectives.
Implement	The overall data governance model.
Monitor	The value of the data and the return on investment.

3.3.8 CSF #8: Proper data integration strategies

'Proper data integration strategies' as a CSF for data governance was the main focus of the second case (STC), so we can argue that this factor plays a critical role in the ability of this particular case (STC) to increase the success of its data governance programme. This CSF consists of different elements, including a unified database for products and services and the overall direction of the data integration. Furthermore, of the elements included, the integration between data quality and the advantages presented to the beneficiaries (employees, customers, suppliers, etc.), such as the employee benefit of having health insurance, is linked with accurate employee data.

This CSF emerged due to current practices in terms of the integration between services and products within the system infrastructure. This is not only a technical issue; the strategy and direction of the data integration are also a problem within the second case (STC). As explained in (Chapter 2, paper 4), STC is facing issues not only with too many databases, but also databases that are too big, which are servicing different products without a proper integration strategy. Hence, the issue has become critical, starting particularly with the customer database, and then the products and services data.

A number of actions are recommended for the employment of this CSF (see Table 3-8), which defining the data integration objectives such as making a unique interface that contains all the customer data across the company's products and services. This would enable the company to change practice by assigning an account manager for each customer, instead of the current approach of assigning an account manager to each customer for each service. In addition, data integrity should be integrated with another entity that the person doing the data entry or the manager needs to consider, thus forcing the beneficiary to insert relevant data to receive a benefit, such as linking the updating of employee data with the insurance offered. Furthermore, implementing an appropriate data infrastructure that fulfils the data integration strategy needs is recommended. The final recommendation is to ensure the value of implementing such infrastructure integration by considering the return on investment in data integration initiatives.

Table 3-8 Summary of the actions recommended for the 'Proper data integration strategies' CSF.

Actions recommended for 'Proper data integration strategies'	
Define	The data integration objectives.
Implement	The data infrastructure to fulfil the data integration needs.
Monitor	The value of data integration.

3.3.9 CSF #9: Effective data monitoring and feedback

Effective data monitoring and feedback consists of data auditing and tracking. In order to have a successful data governance programme, data should be monitored and audited. Our analysis stresses the need to implement a data auditing tool that is included in each implemented system.

CSF #9 accrued as the feedback from the data monitoring team of the second case (STC) is used for the continuous improvement of data quality. This is important for telecommunications enterprises, as they provide and sell data in the form of calls, data, and text messages. These services require continuous monitoring and improvement of the accuracy of the data concerned within the organisation.

In terms of the actions recommended (see Table 3-9), these include defining the key performance indicators (KPIs) for data, as well as having tools and technologies that enable data to be tracked. Auto-auditing activities and improvement should be supported in order to have effective data monitoring and feedback, and monitoring systems should possess proper data. Finally, among the data monitoring tools, data

KPIs should be monitored in order to track the data performance effectively across all systems.

Table 3-9 Summary of the actions recommended for the 'Effective data monitoring and feedback' CSF.

Recommended actions for 'Effective data monitoring and feedback'	
Define	The key performance indicators (KPIs) for data.
Implement	Data monitoring tools within each business system.
Monitor	The data performance against the KPIs for data.

3.4 Summary of the Actions Recommended for the CSFs Identified

From a presentation perspective, we combined the recommended actions into one assessment matrix that considers the outcomes of the CSFs that emerged. Table 3-11 illustrates these actions among the related CSFs and across the three action verbs considered: define, implement, and monitor. This matrix can be used as an action-oriented CSF template to assess the current practice of the data governance programme within an organisation.

In order to obtain the maximum benefit from this assessment tool, it is recommended that the current practice among each of the recommended actions is evaluated by identifying the level of that practice. This can be done using one of the IS maturity assessment levels, such as Control Objectives for Information and Related Technologies (COBIT) (Simonsson, Johnson, & Wijkström, 2007),

Capability Maturity Model Integration (CMMI) (Team, C.P.D, 2000), or the Strategy Alignment Maturity Model (SAMM) (Luftman, 2004). Hence, the scale can be translated into different levels, such as: from 0 to 5; or Initial, Committed, Established, Improved, and Optimised. Table 3-10 offers examples of maturity using the levels in the strategy alignment maturity model (Luftman, 2004) for the CSF 'Clear data processes and procedures' utilising the recommended actions. Each level is represented by a statement that describes the stage of maturity for the data processes and procedures in the context of the data governance programme in an organisation. These statements should be taken further and certain practices developed for each level in order to guide the evaluation process.

Table 3-10 Illustrative examples of the maturity levels for the CSF 'Clear data processes and procedures'

Maturity level	Statement
Initial	The data processes and procedures are not established.
Committed	The capture and retrieval of critical data are defined and there is a commitment to put them in place.
Established	The data capture and retrieval processes are well defined for all organisational data.
Improved	The defined data processes and procedures are embedded in the business systems.
Optimised	The data flow complies with the defined processes and procedures to ensure the optimal use of data assets.

For further research, using one of the maturity models for each of the identified CSFs utilising the recommended actions would enable the assessor to question and visualise the current practice of data governance, as well as to understand the actions that are missing but that should be carried out in order to have an effective data governance programme. However, the three actions are not always undertaken in sequence; in some cases, some actions are implemented but are not well defined. The case of a missing defined action would require more effort from the implementers, as they would be implementing something that had not been fully defined.

Table 3-11 Summary of the actions recommended for the CSFs identified.

CSF	Define	Implement	Monitor
Employee data competencies	The required skills and competencies for dealing with the data for each role and responsibility.	Education and training for employees in 'how to deal with data', as well as 'increasing awareness of the importance of data'.	Employee activities and performance when using data.
Flexible data tools and technologies	Data life cycle requirements to do for integration technical needs.	An appropriate technical architecture to meet integration and the data life cycle needs.	Data life cycle, and data integration.
Clear data processes and procedures	Data capture and retrieval processes for all organisational data.	Data capture and automated validation by embedding them into business systems.	The data flow and data use.
Clear, inclusive data requirements	Data requirement standards and structure template.	An appropriate data requirement standards template.	The communication between parties regarding data requirements specification.
Standardised easy-to-follow data policies	The data regulations, access rights and privileges.	The data policies within the business systems.	The compliance with external and internal data regulations.
Established data roles and responsibilities	The data governance position and data decision rights.	The assignment of data roles to decision areas.	The clarity of data responsibilities.
Focused and tangible data strategies	Data value and objectives.	The overall data governance model.	The value of the data and the return on investment.
Proper data integration strategies	The data integration objectives.	The data infrastructure to fulfil the data integration needs.	The value of data integration.
Effective data monitoring and feedback	The key performance indicators (KPIs) for data.	Data monitoring tools within each business system.	The data performance against the KPIs for data.

3.5 Possible Interconnectedness of the CSFs Identified

The possible interconnectedness of the CSFs identified enables better understanding of the multiplicity of effects of each CSF. Figure 3-2 illustrates a causal map of the possible interconnectedness of the nine CSFs identified as an outcome of the within-and cross-case analysis conducted for each of the case studies. The relationships in the causal map represent logical links between the CSFs that interpret the impact of the presence or absence of one CSF on another. However, where a CSF does not have a direct impact on the other CSFs does not mean that the CSF is of lower importance compared with the others. Potential interconnectedness only shows the possible impact of the presence or absence of one CSF on another.

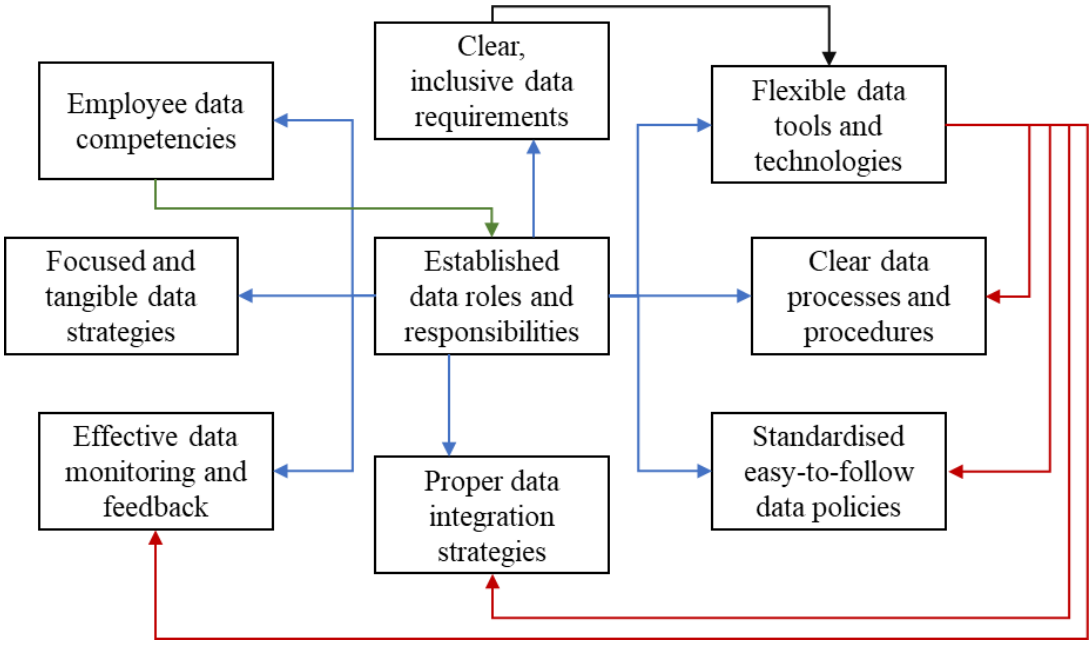

Figure 3-2 Causal map of the possible interconnectedness of the nine CSFs identified.

The causal map of the possible interconnectedness of CSFs provides insights for a data governance committee to understand the sequence of establishing a data governance programme. Although the causal relationships above do not demonstrate the importance of one CSF over another, they illustrate the impact of the presence or absence of one CSF on another. Hence, utilising the causal map (Figure 3-2) suggests a road map for a data governance committee of the sequence of areas or actions that should be considered when establishing a data governance programme.

For example, the causal map shows greater interconnectedness of the CSF ‘Establish data roles and responsibilities with other CSFs’. This suggests commencing a data governance programme by establishing the data governance structure and introducing the various roles and responsibilities, although this should be aligned with identifying the skills required for each. This would encourage other interrelated CSFs, such as ‘Clear data processes and procedures’, to be established in a more effective manner. In addition, the CSF ‘Flexible data tools and technologies’ relies on the CSF ‘Clear inclusive data requirements’, as it drives the tools and technologies required for data governance.

On the other hand, focusing on the CSF 'Clear data processes and procedures' without considering the data roles and responsibilities required might suggest a difficulty or failure in defining and implementing the data processes and procedures, as not establishing data roles and responsibilities leads to uncertainty regarding who should define the data processes and procedures, who should implement them, and who can modify them. The areas of interconnection are

further articulated in Table 3-12, which presents a description of the main areas of impact that the absence/presence of one of the CSFs would have on each of the others.

Table 3-12 Areas of interconnection and possible interconnectedness of the nine CSFs identified.

CSF	Has an impact on	Interconnection
Employee data competencies	Established data roles and responsibilities	Defines the competency requirements of each role.
Flexible data tools and technologies	Clear data processes and procedures	Embeds the data processes and procedures into the systems.
	Standardised easy-to-follow data policies	Embeds the data policies into the systems.
	Proper data integration strategies	Enables the implementation of data integration strategies.
	Effective data monitoring and reporting	Enables the establishment of data monitoring and reporting tools.
Established data roles and responsibilities	All CSFs	Defines who does what and assigns responsibilities for each activity.
Clear inclusive data requirements	Flexible data tools and technologies	Drives the tools and technologies required for data governance.

From Table 3-12, it can be seen that the ‘Employee data competencies’ CSF has a direct impact on ‘Established data roles and responsibilities’, as each data role and responsibility requires certain employee competencies in order to be performed successfully. This has an impact on ‘Established data roles and responsibilities’,

which has a further direct impact on the all of the other CSFs. Therefore, without paying attention to the level of importance of the CSFs reported within our analysis, the actions related to the 'Established data roles and responsibilities' CSF should be considered the first action to be taken in order to guarantee the effective of the other CSFs.

In addition, as shown in Table 3-12, 'Flexible data tools and technologies' has an impact on four of the CSFs due to the governing of data relying heavily on the technologies that are involved. For example, 'Clear data processes and procedures' is impacted along with the current tools and technologies, as these processes and procedures need to be embedded within the business systems as well as along the data life cycle. 'Flexible data tools and technologies' is also interconnected with 'Clear, inclusive data requirements', as this defines the data tools and technologies required within the data governance programme. Finally, in addition to any direct interconnectedness, there is another possible form of connectedness between parts of each CSF with other parts of other CSFs, which can be presented as the interconnectedness between the actions.

3.6 Comparison with the Literature

As stated in the introductory chapter (Chapter 1), there is a limited number of studies that focus on the critical success factors for data governance. However, there are many studies that provide data governance models with a certain focus that can be used to understand part of the actions required for having an effective data governance programme. Therefore, in this section, as recommended by Eisenhardt (1989), an attempt is made to identify and compare similarities and differences, and

what has been reported in the literature as the actions or activities required for conducting a data governance programme, with the CSFs for data governance that emerged from the case studies reported in this research.

In general, the activities reported for data governance programmes in the literature support the CSFs for data governance that were identified from analysing the two case studies. However, the level of the strength of the support varies from one CSF to another. When comparing the frequency count of the reported activities (see Chapter 2, paper 2) for data governance with the level of importance of the recommended actions within the CSFs for data governance, it can be seen that there is a clear difference in focus and importance. For example, the activities related to the area of governance regarding 'Data roles and responsibilities' from the literature point of view have the highest number of reported activities compared with the other areas of governance. On the other hand, from the results of the case studies, the 'Established data roles and responsibilities' CSF is ranked sixth in importance based on the frequency count of the concepts generated during analysis.

From the case studies analysed, the 'Employee data competencies' CSF was the most critical success factor for data governance and includes different recommended actions. From the literature point of view, however, we can see that there is a lack of reported actions related to employee competencies, such as the training, awareness, and skills required for each role and responsibility. Of the publications analysed (see Chapter 2, paper 2) related to data governance activities, none of the publications focused mainly on employee data competencies, although some publications report a single sentence recommending conducting part of these

activities (such as Cheong and Chang, 2007) within the context of explaining the responsibilities of the role of 'Data Steward', stating that "*They manage user group meetings, train and educate data users*" (p. 1005).

Furthermore, the literature reports more 'define'-related activities for data governance, whereas our analysis of the case studies shows a different focus, in particular on the implementation-related actions within each CSF. For example, the implementation of data processes and procedures is shown in the data governance activities model (see Chapter 2, paper 1) as being between medium and low, whereas there is a high volume of actions/ interactions, which indicates that 'implement' actions for data processes and procedures emerged from coding the interviews across both cases. This argument is also applicable within the area of data policies. Hence, it can be argued that, in order to have successful data governance, there is a need to focus on the implementation activities of each of the CSFs identified.

To summarise, comparing the CSFs identified for data governance in this research with findings in the literature cannot be conducted in full due to the limitation of the CSFs reported for data governance in past research. However, the above is an attempt to compare the activities reported for data governance that were identified inductively (see Chapter 2, paper 1) with the actions that are associated with each CSF. We find that there is a high level of similarity in the reported activities, particularly those associated with the 'define' action, as well as in some of the areas of governance, such as 'Data processes and procedures' and 'Data strategies'. However, some of the core elements of data governance found, following analysis

of the cases studies, are infrequently reported in the literature, such as activities related to employee data competencies. This can be taken as an area of research that needs further investigation in a different context in order to demonstrate the actions required related to employee data competencies.

In addition, the CSFs presented can be taken further by categorising them around different possible categories in order to be more valuable and abstracted, such as the categorisation by Lam (2005), in which CSFs are assigned to four areas: Business, Organisation, Technology, and Project, Figure 3-3 illustrates an example of the categorisation of the identified CSFs following (Lam, 2005) categories. Another simple form of categorisation is processes, people, and technology (Fisher, 2006). Categorisation can also indicate the areas of each CSF or the importance level of the CSFs for individuals (c.f. Williams & Ramaprasad, 1996).

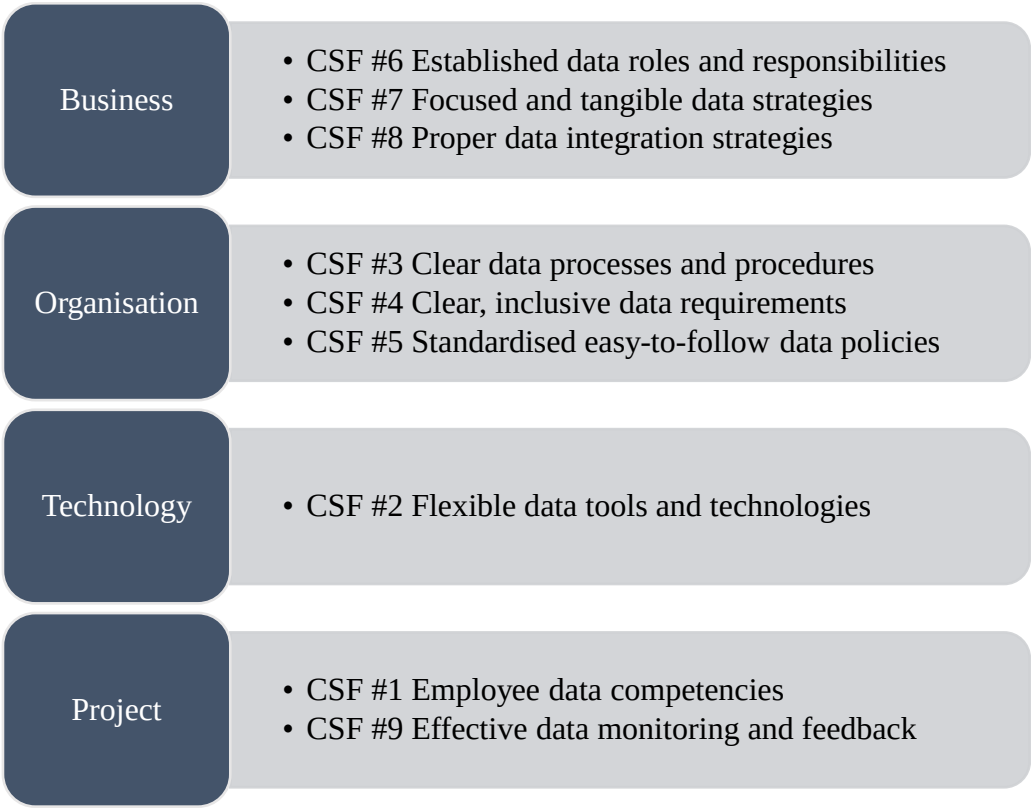

Figure 3-3 Categorisation example of the nine CSFs identified.

3.7 Conclusions and Research Implications

Research in the data governance domain is growing in IS, as is the need for research in this area, as more organisations consider data as a valuable asset (Khatri & Brown, 2010; Otto, 2015). A review of the data governance literature shows that there is a lack of research that explicitly studies the critical success factors for governing data and the activities they include. Nevertheless, there is some research that contributes to our understanding of data governance through modelling (c.f. Khatri & Brown, 2010; Otto, 2011b; Tallon, Ramirez, & Short, 2013). These studies reveal some progress in exploring the activities that are required for

governing data. Therefore, it was decided to conduct this research by following the research strategy of building theory from case study.

This research identified the CSFs for data governance inductively by following a grounded theory approach to theory building. Three main sources to shape the research methodology: 1) The Eisenhardt (1989), approach to building theory from a case study. This approach informed the theory-building road map used in this research; 2) the Rockart (1979) CSF approach. This approach enabled the researcher to follow a data collection procedure within the context of identifying the CSFs for data governance; and 3) the Strauss and Corbin (1990), open, axial, and selective coding approach. This approach provided detailed explanations of how open, axial, and selective coding should be conducted as part of a grounded theory approach. Although Strauss and Corbin's (1990) work formed the researcher's initial of how to operationalise open, axial, and selective coding, there was a need to conduct further research in order to clarify the processes and establish an easy-to-follow framework for conducting these types of coding.

This PhD research study is based on a series of papers (see Chapter 2) that provide an account of the pursuit of a research objective to identify the CSFs for data governance that enable organisations to introduce an effective data governance programme. It started with a chapter that introduced the research road map and research objective and the research motivation, and provided a summary of each of the research components: a literature review, the research methodology, and the findings. The introduction also included descriptions of the topics that are not fully

served within the series of publications included. For example, there is further description of the CSF approach, as well as a definition of data governance.

After the introduction chapter, four papers are organised in the following order. Paper 1 is the literature review chapter and comprehensively reviews the data governance literature (academic and practitioner) for the reader to understand the data governance activities that are reported in the literature. Paper 1 follows a systematic literature review procedure. Paper 2 focuses on the research methodology, specifically on the use of open, axial, and selective (OAS) coding. It investigates the use of OAS coding techniques and the paper is based on reviewing and analysing IS studies that have operationalised the techniques.

The third and fourth papers report the critical success factors for data governance by following an inductive approach aimed at building theory from case study, undertaking a CSF approach to data collection and using OAS coding techniques to conduct the data analysis. Paper 3 identifies the CSFs for data governance that emerged from the first case (Alrajhi Bank). Paper 4 identifies the CSFs for data governance that emerged from the second case (Saudi Telecom Company (STC)).

Drawing on the four papers, this research ends with a discussion and conclusion chapter, in which cross-case analysis is conducted in order to arrive at a final list of CSFs for data governance. This chapter provides a full description of the CSFs identified and presents actions recommended within each CSF. It concludes with an action-oriented CSF template to assess the current state of a data governance programme within an organisation, therefore, understanding the action undertaken

around data governance. The final section presents a comparison of the results of this research with the existing literature on data governance.

The remainder of this conclusions and research implications section highlights the research study contributions for researchers and data governance practitioners, then draws attention to the research implications for theory and practice. Finally, the chapter provides a subsection outlining the research limitations and recommendations for future research.

3.7.1 Research Contributions

This research study offers a number of contributions to both academia and to practice by way of not only what was achieved, but also how this result was accomplished. This section offers a summary of the main contributions form this research (see Table 3-13), as already explained within the conclusions of each of the four papers presented in Chapter 2. In terms of the main contributions to data governance research, this research proposed nine CSFs for data governance. The nine CSFs identified are as follows: 1) Employee data competencies, 2) Flexible data tools and technologies, 3) Clear data processes and procedures, 4) Clear, inclusive data requirements, 5) Standardised easy-to-follow data policies, 6) Established data roles and responsibilities, 7) Focused and tangible data strategies, 8) Proper data integration strategies, and 9) Effective data monitoring and feedback.

In addition to its contributions to data governance research, this study details the possible interconnectedness of the CSFs identified and suggests areas impacted by the presence/absence of each CSF on the others. This helps to prioritise the

implementation of the CSFs in order to have an effective data governance programme.

From the critical success factors identified, this study highlights the recommended actions that should be taken to put the CSFs into practice. These actions are mapped against three action verbs: ‘define’, ‘implement’, and ‘monitor’. The combining of these three actions makes a contribution to practice by producing a data governance assessment matrix that can be considered as a template for assessing the current practice of the data governance programme within an organisation.

Furthermore, this study analysed the data governance literature in order to generate a data governance activities model based on the actions reported in academic as well as practice-oriented publications. This model is intended to help researchers understand the activities involved in conducting a data governance programme, as well as the priorities for each activity.

Finally, this study contributes to IS research by providing a coding framework that supports the decision-making of novice researchers intending to conduct open, axial, and selective coding as part of their qualitative data analysis. This framework was achieved by reviewing and analysing previous IS studies that have operationalised the OAS coding techniques.

Table 3-13 Summary of the contributions of the study.

Contribution to	Contribution
Data governance research	**Identifies nine CSFs for data governance** Identifies nine CSFs for data governance by analysing two case studies.
	Recommended actions for each CSF Identifies a list of recommended actions for each CSF that emerged from analysing the two case studies and maps them against three action verbs: 'define', 'implement', and 'monitor'.
	Interconnectedness of the CSFs identified Depicts the possible interconnectedness of the CSFs identified by showing the areas impacted by the presence/absence of each of the CSFs on the others.
	Data governance activity model Identifies a data governance activity model based on an analysis of the activities reported in the academic and practitioner literature.
IS research methodology	**OAS coding framework** Investigates the usage of open, axial, and selective (OAS) coding techniques by reviewing and analysing IS studies that have operationalised the techniques in order to generate an OAS coding framework.
Practice	**Data governance assessment matrix** Provides a template for assessing the current practice around a data governance programme within an organisation.

3.7.2 Implications for Theory and Practice

The contributions outlined in the previous subsection have implications for future research, as well as the practitioner community. The CSFs identified for data governance should enable researchers to better understand the areas that need to be considered when conducting a data governance programme. However, due to the CSFs having emerged from case studies, it could be argued that there is an opportunity to extend the factors identified by conducting a field study to investigate measures that could shape and evaluate the current practice within each CSF.

In addition, as stated previously, this study contributes to data governance research by highlighting the possible interconnectedness of the CSFs identified. However, in order to have a universal model of CSFs for data governance, it is recommended that further research examine the interconnectedness between the factors in order to better understand the implications of the presence or absence of each one. This would create an opportunity for researchers to test and evaluate propositions that can be made by connecting each of the CSFs. For example, one possible research route would be to test and evaluate whether the 'Clear data processes and procedures' CSF has an impact on the employment of 'Clear, inclusive data requirements'.

Furthermore, in terms of the data governance activity model, as stated in (Chapter 2, paper 1), the activities reported in the scientific publications focus more on 'defining' activities, whereas practice-oriented publications consider the

‘implementation’ and ‘monitoring’ of activities. Therefore, more academic research is needed around the ‘implement’ and ‘monitor’ actions in data governance.

Finally, in terms of practice implications, this study is intended to serve practitioners who are working in different data governance roles by enabling them to better understand the actions that are recommended to be undertaken in order to have an effective data governance programme. The matrix (see Table 3-11) to assess the current data governance practices within an organisation, that emerged from our case study analysis can be taken further and used as an auditing tool for data governance. The assessment matrix includes areas that could be enhanced for a specific organisation or industry that would help to shape the full picture of the data governance programme within that organisation or industry.

3.7.3 Limitations and Future Research

Although this research study endeavoured to achieve the highest levels of objectivity, accuracy and validity, the study is not without limitations. Indeed, despite the best efforts of scholars, research studies will often be constrained by one or more factors, such as time and resources. Any piece of research has limitations. This study has several limitations, which can be addressed by future research.

As this research study was aimed at building theory from case studies, this shows a natural limitation, as the results presented specifically reflect the situation of the selected cases and show only part of the wider picture. The CSFs identified may, however, be useful in providing guidelines for those who want to conduct a data governance programme within a large enterprise. Hence, it is recommended that

further research examine the CSFs identified and the activities involved within each in order to better understand the implementation road map of each factor.

In addition, the possible interconnectedness of the CSFs identified, which articulates the existence of an impact of the presence/absence of every CSF on each, requires further investigation and study in order to specify the interconnectedness of the actions within each CSF. This can be done by further testing the propositions that link the CSFs, as well as the actions included.

Finally, as shown in (Chapter 2, paper 2), this research provides an OAS coding framework and a list of recommendations for novice researchers who intend to use open, axial, and selective coding for analysing content. The framework was built inductively by reviewing how senior IS scholars have used OAS coding. This presents an opportunity for further research to test the framework in different contexts and on different types of data.

References

Abraham C, Boudreau M, Junglas I And Watson R (2013) Enriching our theoretical repertoire: The role of evolutionary psychology in technology acceptance. European Journal of Information Systems, 22(1), 56-75.

Alderson, C 2014, 'Information governance in the NHS: Caldicott 2 and beyond', *Information System*, vol. 19, no. 6, pp. 115-119.

Al-Ruithe, M., Benkhelifa, E., & Hameed, K. (2016). A conceptual framework for designing data governance for cloud computing. Procedia Computer Science, 94, 160-167.

Anderson, DR, Sweeney, DJ, & Williams, TA 2011, *Essentials of statistics for business and economics*, revised, Cengage Learning, Stamford, Connecticut.

Bach, P 2006, 'A 360-degree approach to data governance', *Financial Executive*, vol. 22, no. 6, pp. 54-55.

Bagayogo Ff, Lapointe L And Bassellier G (2014) Enhanced use of IT: A new perspective on post-adoption. Journal of the Association for Information Systems, 15(7), 361-387.

Baskerville R And Pries‐Heje J (2004) Short cycle time systems development. Information Systems Journal, 14(3), 237-264.

Becker, MY 2007, 'Information governance in NHS's NPfIT: A case for policy specification', *International Journal of Medical Informatics*, vol. 76, no. 5-6, pp. 432-437. Available from: http://0-dx.doi.org.library.ucc.ie/10.1016/j.ijmedinf.2006.09.008.

Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The case research strategy in studies of information systems. *MIS Quarterly*, , 369-386.

Benbunan-Fich, R. And Benbunan, A., 2007. Understanding user behavior with new mobile applications. The Journal of Strategic Information Systems, 16(4), pp. 393-412.

Berente N And Yoo Y (2012) Institutional contradictions and loose coupling: Post Implementation of NASA's enterprise information system. Information Systems Research, 23(2), 376-396.

Berente N, Gal U And Yoo Y (2010) Dressage, control, and enterprise systems: The case of NASA's full cost initiative. European Journal of Information Systems, 19(1), 21-34.

Berente N, Hansen S, Pike Jc And Bateman Pj (2011) Arguing the value of virtual worlds: Patterns of discursive sensemaking of an innovative technology. MIS Quarterly, 35(3), 685-709.

Berg, B. L. (2001). Methods for the social sciences Pearson Education Inc, United States of America.

Bhattacherjee, A. (2012). Social science research: principles, methods, and practices (2nd ed.). Tampa, FL: CreateSpace Independent Publishing Platform.

Birks Df, Fernandez W, Levina N And Nasirin S (2013) Grounded theory method in information systems research: Its nature, diversity and opportunities. European Journal of Information Systems, 22(1), 1-8.

Blair, BT 2010, 'Governance for protecting information in the cloud', *Information Management*, vol. 44, no. 5, pp. HT1-HT4. Available from: http://crawl.prod.proquest.com.s3.amazonaws.com/fpcache/a11e60986ad8e893b3544cc652658610.pdf?AWSAccessKeyId=AKIAJF7V7KNV2KKY2NUQ&Expires=1476691882&Signature=n5rAaY1La38h%2Fv5oL%2FQaqna%2FXZA%3D.

Bowen, R & Smith, AR 2014, 'Developing an enterprisewide data strategy', *Healthcare Financial Management*, vol. 68, no. 4, pp. 86-89.

Bryman, A., 1984. The Debate about Quantitative and Qualitative Research: A Question of Method or Epistemology? The British journal of sociology, 35(1), pp. 75-92.

Buchwald, A., Urbach, N., & Ahlemann, F. (2014). Business value through controlled IT: toward an integrated model of IT governance success and its impact. Journal of Information Technology, 29(2), 128-147. Retrieved from https://link.springer.com/article/10.1057/jit.2014.3

Butler T (2011) Compliance with institutional imperatives on environmental sustainability: Building theory on the role of green IS. The Journal of Strategic Information Systems, 20(1), 6-26.

Butler T And Murphy C (2008) An exploratory study on IS capabilities and assets in a small-to-medium software enterprise. Journal of Information Technology, 23(4), 330-344.

Butler, T., & Fitzgerald, B. (1999). Unpacking the systems development process: an empirical application of the CSF concept in a research context. The Journal of Strategic Information Systems, 8(4), 351-371. doi:10.1016/S0963-8687(00)00027-5

Caralli, R.A., Stevens, J.F., Willke, B.J., & Wilson, W.R. (2004) The critical success factor method: establishing a foundation for enterprise security management. Retrieved from http://www.dtic.mil/get-tr-doc/pdf?AD=ADA443742

Carley, K 1993, 'Coding choices for textual analysis: A comparison of content analysis and map analysis', *Sociological Methodology*, vol. 23, no. 75-126.

CDI Institute 2006, *Corporate data governance best practices: 2006-07 scorecards for data governance in the global 5000*, Burlingame, CA: CDI Institute.

Chakraborty S, Sarker S And Sarker, S (2010) An exploration into the process of requirements elicitation: A grounded approach. Journal of the Association for Information Systems, 11(4), 212-249.

Chan Cml, Hackney R, Pan Sl And Chou T (2011) Managing e-government system implementation: A resource enactment perspective. European Journal of Information Systems, 20(5), 529-541.

Chang Clh, Chen V, Klein G And Jiang Jj (2011) Information system personnel career anchor changes leading to career changes. European Journal of Information Systems, 20(1), 103-117.

Cheong, LK & Chang, V (2007) 'The need for data governance: A case study', *Proceedings of the 18th Australasian conference on information systems*, ACIS 2007 Proceedings, pp. 999-1008.

Clemmensen T (2012) Usability problem identification in culturally diverse settings. Information Systems Journal, 22(2), 151-175.

Cohen, R (2006) What's in a name? Data governance roles, responsibilities and results factors, Information Management. Available from: http://www.information-management.com/news/columns/-1057220-1.html. [8 June 2006].

Cousins, K. (2016). Health IT legislation in the united states: Guidelines for IS researchers. Communications of the Association for Information Systems, 39(1), 17.

DAMA International 2009, *The DAMA guide to the data management body of knowledge*, Denville, New Jersey, Technics Publications.

Darke, P., Shanks, G., & Broadbent, M. (1998). Successfully completing case study research: Combining rigour, relevance and pragmatism. *Information Systems Journal*, *8*(4), 273-289.

Day Jm, Junglas I And Silva L (2009) Information flow impediments in disaster relief supply chains. Journal of the Association for Information Systems, 10(8), 637-660.

Dember, M 2006, '7 stages for effective data governance', *Architecture & Governance Magazine*, vol. 2, no. 4, pp. 1-3.

Dezdar, S., & Sulaiman, A. (2009). Successful enterprise resource planning implementation: Taxonomy of critical factors. *Industr Mngmnt & Data Systems, 109*(8), 1037-1052. doi:10.1108/02635570910991283

Donaldson, A & Walker, P 2004, 'Information governance—a view from the NHS', *International Journal of Medical Informatics*, vol. 73, no. 3, pp. 281-284. Available from: http://dx.doi.org/10.1016/j.ijmedinf.2003.11.009.

Dyché, J 2007, *A data governance manifesto: Designing and deploying sustainable data governance*, San Mateo, California: Siperian, Inc.

Eaves, Y. D. (2001). A synthesis technique for grounded theory data analysis. Journal of Advanced Nursing, 35(5), 654-663.

Economist Intelligence Unit 2008, 'The future of enterprise information governance', *The Economist*. Available from: http://viewswire.eiu.com/report_dl.asp?mode=fi&fi=1643883749.PDF.

Eisenhardt, K.M. (1989). Building Theories from Case Study Research. The Academy of Management Review, 14(4), 532-550. doi:10.5465/AMR.1989.4308385

Elliott, TE, Holmes, JH, Davidson, AJ, Chance, L, Nelson, AF, & Steiner, JF 2013, 'Data warehouse governance programs in healthcare settings: A literature review and a call to action', *EGEMs (Generating Evidence & Methods to Improve Patient Outcomes)*, vol. 1, no. 1, pp. 15.

Esterberg, K.G. (2002). Qualitative methods in social research. Boston, MA: McGraw Hill.

Feldman Ss And Horan Ta (2011) The dynamics of information collaboration: A case study of blended it value propositions for health information exchange in disability determination. Journal of the Association for Information Systems, 12(2), 189-207.

Feller J, Finnegan P, Fitzgerald B And Hayes J (2008) From peer production to productization: A study of socially enabled business exchanges in open source service networks. Information Systems Research, 19(4), 475-493.

Feller, J, Finnegan P, Hayes J And O'reilly P (2012) 'Orchestrating' sustainable crowdsourcing: A characterisation of solver brokerages. The Journal of Strategic Information Systems, 21(3), 216-232.

Finney, S & Corbett, M 2007, 'ERP implementation: A compilation and analysis of critical success factors', *Business Process Management Journal*, vol. 13, no. 3, pp. 329-347. Available from: http://0-dx.doi.org.library.ucc.ie/10.1108/14637150710752272.

Fisher, T 2006, 'Data monitoring: Add controls to your data governance and compliance programs', *Business Intelligence Journal*, vol. 11, pp. 51-57.

Fu, X, Wojak, A, Neagu, D, Ridley, M., & Travis, K 2011, 'Data governance in predictive toxicology: A review', *Journal of Cheminformatics*, vol. 3, no. 1, pp. 1-16.

Furneaux B And Wade Mr (2011) An exploration of organizational level information systems discontinuance intentions. MIS Quarterly, 35(3) , 573-598.

Galal G (2001) From contexts to constructs: The use of grounded theory in operationalising contingent process models. European Journal of Information Systems, 10(1), 2-14.

Gillies, A & Howard, J 2005, 'An international comparison of information in adverse events', *International Journal of Health Care Quality Assurance*, vol. 18, no. 5, pp. 343-352.

Given, L. M. (2008). The sage encyclopedia of qualitative research methods Sage Publications.

Glaser, B. G. (1992). Emergence vs forcing: Basics of Grounded Theory Analysis. Mill Valley, CA: Sociology Press.

Glaser, B., & Strauss, A. (1967). The discovery grounded theory: Strategies for qualitative inquiry. *Aldin, Chicago,*

Glaser, BG 1992, *Basics of grounded theory analysis: Emergence vs. forcing*, Mill Valley, California, Sociology Press.

Gleasure R (2015) Resistance to crowdfunding among entrepreneurs: An impression management perspective. The Journal of Strategic Information Systems, 24(4), 219-233.

Goo, J., Kishore, R., Rao, H.R. And Nam, K. (2009) The role of service level agreements in relational management of information technology outsourcing: an empirical study. MIS Quarterly, , pp. 119-145.

Goode S And Gregor S (2009) Rethinking organisational size in IS research: Meaning, measurement and redevelopment. European Journal of Information Systems, 18(1), 4-25.

Goulielmos M (2004) Systems development approach: Transcending methodology. Information Systems Journal, 14(4), 363-386.

Grahlmann, K.R., Helms, R.W., Hilhorst, C., Brinkkemper, S. And Van Amerongen, S. (2012) Reviewing enterprise content management: A functional framework. European Journal of Information Systems, 21(3), pp. 268-286.

Guetat, SBA & Dakhli, SBD (2015) 'The architecture facet of information governance: The case of urbanized information systems', *Procedia Computer Science*, vol. 64,

pp. 1088-1098. Available from: http://0-dx.doi.org.library.ucc.ie/10.1016/j.procs.2015.08.564.

Guynes, C.S., & Vanecek, M.T. (1996). Critical success factors in data management. Information & Management, 30(4), 201-209. doi:10.1016/0378-7206(95)00053-4

Hackney R, Jones S And Lösch A (2007) Towards an e-government efficiency agenda: The impact of information and communication behaviour on e-reverse auctions in public sector procurement. European Journal of Information Systems, 16(2), 178-191.

Hagmann, J. 2013, 'Information governance – beyond the buzznull', *Records Management* Journal, vol. 23, no.3, pp. 228-240.

Hassan, S., & Chindamo, P. (2017). Governance in practice: Effective data governance: From strategy through to implementation. Governance Directions, 69(4), 207.

Heath, H., & Cowley, S. (2004). Developing a grounded theory approach: A comparison of glaser and strauss. International Journal of Nursing Studies, 41(2), 141-150.

Holmström J And Sawyer S (2011) Requirements engineering blinders: Exploring information systems developers' black-boxing of the emergent character of requirements. European Journal of Information Systems, 20(1), 34-47.

Holt, V., Ramage, M., Kear, K., & Heap, N. (2015). The usage of best practices and procedures in the database community. Information Systems, 49, 163-181.

Horne, N.W. (1995). Information as an asset—The board agenda. Computer Audit Update, 1995(9), 5-11. doi:10.1016/0960-2593(95)90246-5

Huang J, Makoju E, Newell S And Galliers Rd (2003) Opportunities to learn from 'failure' with electronic commerce: a case study of electronic banking. Journal of Information Technology, 18(1), 17-26.

Huang Jc, Newell S And Pan Sl (2001) The process of global knowledge integration: A case study of a multinational investment bank's Y2K program. European Journal of Information Systems, 10(3), 161-174.

Huang P, Pan Sl And Ouyang Th (2014) Developing information processing capability for operational agility: Implications from a chinese manufacturer. European Journal of Information Systems, 23(4), 462-480.

Hutchinson, C & Sharples, C (2006) 'Information governance: Practical implications for record-keeping', *Nursing Standard*, vol. 20, no.36, pp. 59-64.

IBM 2007, *The IBM data governance council maturity model: Building a roadmap for effective data governance*, Somers, New York: IBM Corporation.

Informatica 2013, *Data governance implementation: Measuring the success, compliance and effectiveness of enterprise data*, Redwood, California: Informatica.

Information Builders 2011, *Seven steps to effective data governance: A white paper*, New York, New York: Information Builders.

Information Builders. (2013), Driving Better Business Performance With a Practical Data Strategy, [White paper].

Jenkin Ta And Chan Ye (2010) IS project alignment–a process perspective. Journal of Information Technology, 25(1), 35-55.

Johnson, R.B. and Onwuegbuzie, A.J., 2004. Mixed methods research: A research paradigm whose time has come. Educational researcher, 33(7), pp. 14-26.

Kamioka, T., Luo, X., & Tapanainen, T. (2016). An empirical investigation of data governance: The role of accountabilities. The 20th Pacific Asia Conference on Information Systems (PACIS),

Kane Gc And Labianca G (2011) IS avoidance in health-care groups: A multilevel investigation. Information Systems Research, 22(3), 504-522.

Kaplan, B. and Duchon, D., 1988. Combining qualitative and quantitative methods in information systems research: a case study. MIS quarterly, , pp. 571-586.

Karoui M, Dudezert A And Leidner De (2015) Strategies and symbolism in the adoption of organizational social networking systems. The Journal of Strategic Information Systems, 24(1), 15-32.

Keil M, Im Gp, And Mähring M (2007) Reporting bad news on software projects: The effects of culturally constituted views of face‐saving. Information Systems Journal, 17(1), 59-87.

Kersten, S 2013, 'Moving the needle toward a data-driven health care system: Optimizing the EHR through information governance', *Journal of Health Care Compliance*, vol. 15, no. 3, pp. 45-62.

Khatcherian, S & Jefferson, G 2009, 'Information governance: Challenges for financial services firms operating in the information age', *Convergence*, vol. 5, no. 1, pp. 80-89.

Khatri, V., & Brown, C.V. (2010). Designing data governance. Communications of the ACM, 53(1), 148-152. doi:10.1145/1629175.1629210

Kirsch Lj (2004) Deploying common systems globally: The dynamics of control. Information Systems Research, 15(4), 374-395.

Kirsch Lj And Haney Mh (2006) Requirements determination for common systems: Turning a global vision into a local reality. The Journal of Strategic Information Systems, 15(2), 79-104.

Kirsch Ls (1997) Portfolios of control modes and IS project management. Information Systems Research, 8(3), 215-239.

Kock N (2001) Asynchronous and distributed process improvement: The role of collaborative technologies. Information Systems Journal, 11(2), 87-110.

Koh, S.C.L., Gunasekaran, A., & Goodman, T. (2011). Drivers, barriers and critical success factors for ERPII implementation in supply chains: A critical analysis. The Journal of Strategic Information Systems, 20(4), 385-402. doi: 10.1016/j.jsis.2011.07.001

Kooper, MN, Maes, R, & Lindgreen, EEOR 2011, 'On the governance of information: Introducing a new concept of governance to support the management of information', *International Journal of Information Management*, vol. 31, no. 3, pp. 195-200. Available from: http://dx.doi.org/10.1016/j.ijinfomgt.2010.05.009.

Lajara, TT & Maçada, ACG (2013) 'Information governance framework: The defense manufacturing case study', *Proceedings of the nineteenth Americas conference on information systems*. Available from: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1704&context=amcis2013. [15-17 August 2013].

Lam, W. (2005). Investigating success factors in enterprise application integration: a case-driven analysis. European Journal of Information Systems, 14(2), 175-187. Retrieved from https://link.springer.com/article/10.1057/palgrave.ejis.3000530

Larkin, BS 2008, 'Increasing information integrity: Cultural impacts of changing the way we manage data', *International Journal of Organisation Theory & Behavior (PrAcademics Press)*, vol. 11, no. 4, pp. 558-578.

Lederman R And Johnston Rb (2011) Decision support or support for situated choice: Lessons for system design from effective manual systems. European Journal of Information Systems, 20(5), 510-528.

Leonardi Pm (2014) Social media, knowledge sharing, and innovation: Toward a theory of communication visibility. Information Systems Research, 25(4), 796-816.

Leonardi, P.M. (2013) When does technology use enable network change in organizations? A comparative study of feature use and shared affordances. MIS Quarterly, 37(3) , 749-775.

Levina, N. And Ross, J.W. (2003) From the vendor's perspective: exploring the value proposition in information technology outsourcing. MIS Quarterly, , pp. 331-364.

Lindgren R, Henfridsson O And Schultze U (2004) Design principles for competence management systems: A synthesis of an action research study. MIS Quarterly, 28(3), 435-472.

Lomas, E 2010, 'Information governance: Information security and access within a UK context', *Records Management Journal*, vol. 20, no. 2, pp. 182-198.

Loshin, D 2013, *Data governance for master data management and beyond: A white paper*, Cary, North Carolina: SAS Institute Inc. Available from: http://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/data-governance-for-MDM-and-beyond-105979.pdf.

Lucas, A 2010, 'Corporate data quality management: From theory to practice', *paper presented at 5th Iberian Conference on Information Systems and Technologies (CISTI)*, Santiago de Compostela, Spain. Available from: http://hdl.handle.net/10400.5/2089. Available from: http://hdl.handle.net/10400.5/2089. [16-19 June 2010].

Luftman, J. N. (2004). Managing the Information Technology Resources. New Jersey: Pearson Prentice Hall. McAfee,

Maldonado E (2010) The process of introducing FLOSS in the public administration: The case of Venezuela. Journal of the Association for Information Systems, 11(11), 756-783.

Martin, JL, Simons, L, Craven, MP, & Betton, V 2014, 'Towards a framework for the appraisal of digital products for mental health', *International Journal of Integrated Care*, vol. 14, no. 8, pp. 1-2. Available from: http://doi.org/10.5334/ijic.1814.

Matavire R And Brown I (2013) Profiling grounded theory approaches in information systems research. European Journal of Information Systems, 22(1), 119-129.

Matsuo M, Wong Cwy And Lai K (2008) Experience-based learning of Japanese IT professionals: A qualitative research. The Journal of Strategic Information Systems, 17(3), 202-213.

Mattarelli E, Bertolotti F And Macrì Dm (2013) The use of ethnography and grounded theory in the development of a management information system. European Journal of Information Systems, 22(1), 26-44.

Meyers, C 2014, ‘How data management and governance can enable successful self-service BI’, *BI Training Solutions: As Close as Your*, vol. 19, no. 4, pp. 23.

Miles, M. B. (1979). Qualitative data as an attractive nuisance: The problem of analysis. *Administrative Science Quarterly, 24*(4, Qualitative Methodology), 590-601.

Miranda, S. M., Kim, I., & Summers, J. D. (2015). Jamming with social media: How cognitive structuring of organizing vision facets affects IT innovation diffusion. MIS Quarterly, 39(3)

Moghe, P 2009, ‘Controlling risk with a data governance framework’, *Bank Accounting & Finance*, pp. 49-51.

Morgan L, Feller J And Finnegan P (2013) Exploring value networks: Theorising the creation and capture of value with open source software. European Journal of Information Systems, 22(5), 569-588.

Myers, M.D., 1997. Qualitative research in information systems. Management Information Systems Quarterly, 21, pp. 241-242.

Myers, M. (2009) Qualitative Research in Business and Management, Sage, UK.

Nagle, T and Sammon, D, (2017). "The Data Value Map: A Framework For Developing Shared Understanding On Data Initiatives". In Proceedings of the 25th European Conference on Information Systems (ECIS), Guimarães, Portugal, June 5-10, 2017 (pp. 1439-1452). ISBN 978-989-20-7655-3 Research Papers. https://aisel.aisnet.org/ecis2017_rp/93

Nwolie, IM 2011, ‘Information governance – what does it mean for you?’, *Practice Nurse*, vol. 41, no. 1, pp. 1-8.

Olsson Hh, Conchúir Eó, Ågerfalk Pj And Fitzgerald B (2008) Two-stage offshoring: An investigation of the irish bridge. MIS Quarterly, 32(2) , 257-279.

Oppenheim, C., Stenson, J., & Wilson, R.M.S. (2003). Studies on Information as an Asset I: Definitions. Journal of Information Science, 29(3), 159-166. doi:10.1177/01655515030293003

Oracle 2011, *Enterprise information management: Best practices in data governance*, Redwood Shores, California: Oracle.

O'reilly P And Finnegan P (2010) Intermediaries in inter-organisational networks: Building a theory of electronic marketplace performance. European Journal of Information Systems, 19(4), 462-480.

Otto, B 2011a, ‘A morphology of the organisation of data governance’, *paper presented at ECIS 2011 proceedings*. Available from: http://aisel.aisnet.org/ecis2011. [6 October 2011].

Otto, B 2011b, ‘Organizing data governance: Findings from the telecommunications industry and consequences for large service providers’, *Communications of the Association for Information Systems*, vol. 29, no. 3, pp. 45-66. Available from: http://aisel.aisnet.org/cais/vol29/iss1/3.

Otto, B 2011c, ‘Data governance’, *Business & Information Systems Engineering*, vol. 3, no. 4, pp. 241-244. Available from: http://dx.doi.org.library.capella.edu/10.1007/s12599-011-0162-8.

Otto, B 2012, ‘Managing the business benefits of product data management: The case of Festo’, *Journal of Enterprise Information Management*, vol. 25, no. 3, pp. 272-297. Available from: http://dx.doi.org/10.1108/17410391211224426.

Otto, B, Wende, K, Schmidt, A, & Osl, P 2007, 'Towards a framework for corporate data quality management', *Proceedings of the 18th Australasian conference on information systems*, ACIS 2007 Proceedings, pp. 916-926.

Otto, B. (2015). Quality and value of the data resource in large enterprises. Information Systems Management, 32(3), 234-251. doi:10.1080/10580530.2015.1044344

Palczewska, A, Fu, X, Trundle, P, Yang, L, Neagu, D, Ridley, M, & Travis, K 2013, 'Towards model governance in predictive toxicology', *International Journal of Information Management*, vol. 33, no. 3, pp. 567-582. Available from: http://dx.doi.org/10.1016/j.ijinfomgt.2013.02.005.

Palka W, Pousttchi K And Wiedemann Dg (2009) Mobile word-of-mouth–A grounded theory of mobile viral marketing. Journal of Information Technology, 24(2), 172-185.

Panian, Z. (2010). Some practical experiences in data governance. World Academy of Science, Engineering and Technology Management, 62, 468-475.

Pauleen Dj (2003) An inductively derived model of leader-initiated relationship building with virtual team members. Journal of Management Information Systems, 20(3), 227-256.

Petrini M And Pozzebon M (2009) Managing sustainability with the support of business intelligence: Integrating socio-environmental indicators and organisational context. The Journal of Strategic Information Systems, 18(4), 178-191.

Pierce, E, Dismute, WS, & Yonke, CL 2008, Industry report: The state of information and data governance: Understanding how organisations govern their information and data assets, Baltimore, Maryland: International Association for Information and Data Quality.

Porter, B. A. (2009). The audit trinity: The key to securing corporate accountability. Managerial Auditing Journal, 24(2), 156-182.

Ramesh B, Cao L And Baskerville R (2010) Agile requirements engineering practices and challenges: An empirical study. Information Systems Journal, 20(5), 449-480.

Ransbotham S And Mitra S (2009) Choice and chance: A conceptual model of paths to information security compromise. Information Systems Research, 20(1), 121-139.

Rasouli, M. R., Eshuis, R., Trienekens, J. J., Kusters, R. J., & Grefen, P. W. J. (2016). Information governance as a dynamic capability in service oriented business networking. Working Conference on Virtual Enterprises, 457-468.

Rau, K. G. (2004). Effective governance of it: Design objectives, roles, and relationships. Information Systems Management, 21(4), 35-42. doi:10.1201/1078/44705.21.4.20040901/84185.4

Redman, T. C. (2013). Data's credibility problem. Harvard Business Review, 91(12), 84-88.

Reeves, MG & Bowen, R 2013, 'Developing a data governance model in health care', Healthcare Financial Management, vol. 67, no. 2, pp. 82-86.

Rickards, RC & Ritsert, R 2012, 'Data governance challenges facing controllers', *International Journal of Business, Accounting, & Finance*, vol. 6, no. 1, pp. 25-42.

Rifaie, M, Alhajj, R, & Ridley, MJ 2009, 'Data governance strategy: A key issue in building enterprise data warehouse', *paper presented at the eleventh international*

conference on information integration and web-based applications and services. [14-16 December 2009].

Rockart, J.F. (1979). Chief executives define their own data needs. Harvard business review, 57(2), 81-93. Retrieved from http://europepmc.org/abstract/med/10297607.

Rosenbaum, S 2010, 'Data governance and stewardship: Designing data stewardship entities and advancing data access', *Health Services Research*, vol. 45, no. 5p2, pp. 1442-1455.

Rowe, F. (2014) "What literature review is not: diversity, boundaries and recommendations", European Journal of Information Systems, 23(3), pp.241-255.

Russom, P 2008, Data governance strategies: Helping your organization comply, transform, *and integrate*, Renton, Washington: TDWI.

Ryan G And Valverde M (2006) Waiting in line for online services: A qualitative study of the user's perspective. Information Systems Journal, 16(2), 181-211.

Sammon, D., & Adam, F. (2008). Justifying an ERP investment: critical success factors for transformation investments. ICIS 2008 Proceedings, 60.

Scott Je (2000) Facilitating interorganizational learning with information technology. Journal of Management Information Systems, 17(2), 81-113.

SEIDEL S And URQUHART C (2013) On emergence and forcing in information systems grounded theory studies: The case of strauss and corbin. Journal of Information Technology, 28(3), 237-260.

Shah, M.H., Braganza, A., & Morabito, V. (2007). A survey of critical success factors in e-Banking: an organisational perspective. European Journal of Information Systems, 16(4), 511-524. Retrieved from https://link.springer.com/article/10.1057/palgrave.ejis.3000693

Shang, S. And Seddon, P.B., 2002. Assessing and managing the benefits of enterprise systems: the business manager's perspective. Information systems journal, 12(4), pp. 271-299.

Shaw-Taylor, Y 2014, 'Making quality improvement programs more effective', *International J Health Care QA*, vol. 27, no. 4, pp. 264-270.

Sheridan, PT & Watzlaf, V 2016, 'Adaptive leadership in information governance', *Journal of AHIMA*, vol. 87, no. 2, pp. 20-25.

Silic, M & Back, A 2013, 'Factors impacting information governance in the mobile device dual-use contextnull', *Records Management Journal*, vol. 23, no. 2, pp. 73-89.

Simonsson, M., Johnson, P., & Wijkström, H. (2007). (2007). Model-based IT governance maturity assessments with COBIT. Paper presented at the Ecis, 1276-1287.

Smolander K, Rossi M And Purao S (2008) Software architectures: Blueprint, literature, language or decision? European Journal of Information Systems, 17(6), 575-588.

Strauss, A., & Corbin, J. (1990). Basics of qualitative research: Grounded theory procedures and techniques. Thousand Oaks, CA: SAGE Publications, Inc.

Strauss, A.L. And Corbin, J. (1998). Basics of Qualitative Research: Techniques and procedures for developing grounded theory, 2nd edn, London: Sage

Strauss, A.L. And Corbin, J. (2008). Basics of Qualitative Research, Thousand Oaks, CA: Sage

Strong Dm And Volkoff O (2010) Understanding Organization—Enterprise system fit: A path to theorizing the information technology artifact. MIS Quarterly, 34(4) , 731-756.

Strong Dm, Johnson Sa, Bengisu T, Trudel J, And Volkoff O (2014) A theory of organization-EHR affordance actualization. Journal of the Association for Information Systems, 15(2), 53-85.

Suer, M & Nolan, R 12 January 2015, *Using COBIT 5 to deliver information and data governance*, COBIT Focus. Available from: http://www.isaca.org/cobit/focus/pages/using-cobit-5-to-deliver-information-and-data-governance.aspx. [7 December 2016].

Susman, G. I., And Evered, R. D. (1978). An assessment of the scientific merits of action research. Administrative science quarterly, 582-603.

Sutton, R.I. and Staw, B.M., 1995. What Theory is Not. Administrative Science Quarterly, 40(3), pp. 371-384.

Tallon, P.P., Ramirez, R.V., & Short, J.E. (2013). The Information Artifact in IT Governance: Toward a Theory of Information Governance. Journal of Management Information Systems, 30(3), 141-178. doi:10.2753/MIS0742-1222300306

Tallon, PP, Short, JE, & Harkins, MW (2013) 'The evolution of information governance at intel', *MIS Quarterly Executive*, vol. 12, no. 4, pp. 189-198.

Tan B, Pan Sl, Lu X And Huang L (2015) The role of IS capabilities in the development of multi-sided platforms: The digital ecosystem strategy of Alibaba. com. Journal of the Association for Information Systems, 16(4), 248-280.

Tan, W.G., Catersteel, A., & Toleman, M. (2009). Implementing it service management: A case study focussing on critical success factors. Journal of Computer Information Systems, 50(2),1-12. doi:10.1080/08874417.2009.11645379

Team, C.P.D. (2000) CMMISM for Systems Engineering/Software Engineering, Version 1.02, Staged Representation, CMU/SEI-2000-TR-018, ESC-TR-2000-018, Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, 1-647

The Data Warehousing Institute 2010, *TDWI data governance fundamentals*, TDWI, Renton, Washington.

Thomas, G 2006, *The DGI data governance framework*, The Data Governance Institute. Available from: http://www.datagovernance.com/the-dgi-framework/. [7 December 2016].

Urquhart C And Fernandez W (2013) Using grounded theory method in information systems: The researcher as blank slate and other myths. Journal of Information Technology, 28(3), 224-236.

Urquhart C, Lehmann H And Myers Md (2010) Putting the 'theory' back into grounded theory: Guidelines for grounded theory studies in information systems. Information Systems Journal, 20(4), 357-381.

Vaast E (2007) Danger is in the eye of the beholders: Social representations of information systems security in healthcare. The Journal of Strategic Information Systems, 16(2), 130-152.

Vaast E And Walsham G (2013) Grounded theorizing for electronically mediated social contexts. European Journal of Information Systems, 22(1), 9-25.

Vannoy Sa And Salam A (2010). Managerial interpretations of the role of information systems in competitive actions and firm performance: A grounded theory investigation. Information Systems Research, 21(3), 496-515.

Vayghan, J. A., Garfinkle, S. M., Walenta, C., Healy, D. C., & Valentin, Z. (2007). The internal information transformation of IBM. IBM Systems Journal, 46(4), 669-683. doi:10.1147/sj.464.0669

Walsh I, Kefi H And Baskerville R (2010) Managing culture creep: Toward a strategic model of user IT culture. The Journal of Strategic Information Systems, 19(4), 257-280.

Walsham, G. (2006). Doing interpretive research. European journal of information systems, 15(3), 320-330. Retrieved from https://link.springer.com/article/10.1057/palgrave.ejis.3000589

Wang, RW & Strong, DM (1996) 'Beyond accuracy: What data quality means to data consumers', *Journal of Management Information Systems*, vol. 12, no. 4, pp. 5-33.

Wang, RY (1998) 'A product perspective on total data quality management', *Communications of the ACM*, vol. 41, no. 2, pp. 58-65.

Watson, HJ, Fuller, C, & Ariyachandra, T (2004) 'Data warehouse governance: Best practices at blue cross and blue shield of north carolina', *Decision Support Systems*, vol. 38, no. 3, pp. 435-450. Available from: http://dx.doi.org/10.1016/j.dss.2003.06.001.

Webb B And Mallon B (2007) A method to bridge the gap between breadth and depth in IS narrative analysis. Journal of the Association for Information Systems, 8(7), 368-381.

Weber, K., Otto, B., & Österle, H. (2009). One Size Does Not Fit All--A Contingency Approach to Data Governance. Journal of Data and Information Quality (JDIQ), 1(1). doi:10.1145/1515693.1515696

Webster J And Watson Rt (2002) Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly, 26(2), xiii-xxiii.

Weill, P & Ross, JW 2004, *IT governance: How top performers manage IT decision rights for superior results*, Harvard Business School Publishing, Boston, Massachusetts.

Weller, A 2008, 'Data governance: Supporting datacentric risk management', *Journal of Securities Operations & Custody*, vol. 1, no. 3, pp. 250-262.

Wende, K & Otto, B (2007) 'A contingency approach to data governance'. In: *12th International Conference on Information Quality (ICIQ-07)*, pp. 163-176. Available from: https://www.alexandria.unisg.ch/213308/. [10 November 2007].

Wende, K (2007) 'A model for data governance–Organising accountabilities for data quality management', *paper presented at the 18th Australasian Conference on Information Systems*, ACIS 2007 Proceedings. Available from: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1079&context=acis2007. [5-7 December 2007].

Williams Ck And Karahanna E (2013) Causal explanation in the coordinating process: A critical realist case study of federated IT governance structures. MIS Quarterly, 37(3), 933-964.

Williams, J., & Ramaprasad, A. (1996). A taxonomy of critical success factors. European Journal of Information Systems, 5(4), 250-260.

Wolfswinkel Jf, Furtmueller E And Wilderom Cp (2013) Using grounded theory as a method for rigorously reviewing literature. European Journal of Information Systems, 22(1), 45-55.

Wood, D 2013, 'Information governance', Simpl Group Ltd, Auckland.

Work B (2002) Patterns of software quality management in TickIT certified firms. European Journal of Information Systems, 11(1), 61-73.

Xu P And Ramesh B (2007) Software process tailoring: An empirical investigation. Journal of Management Information Systems, 24(2), 293-328.

Yin, R. K. (2009). Case study research: Design and Methods. SAGE publications. Thousand oaks.

Zahedi Fm And Bansal G (2011) Cultural signifiers of web site images. Journal of Management Information Systems, 28(1), 147-200.