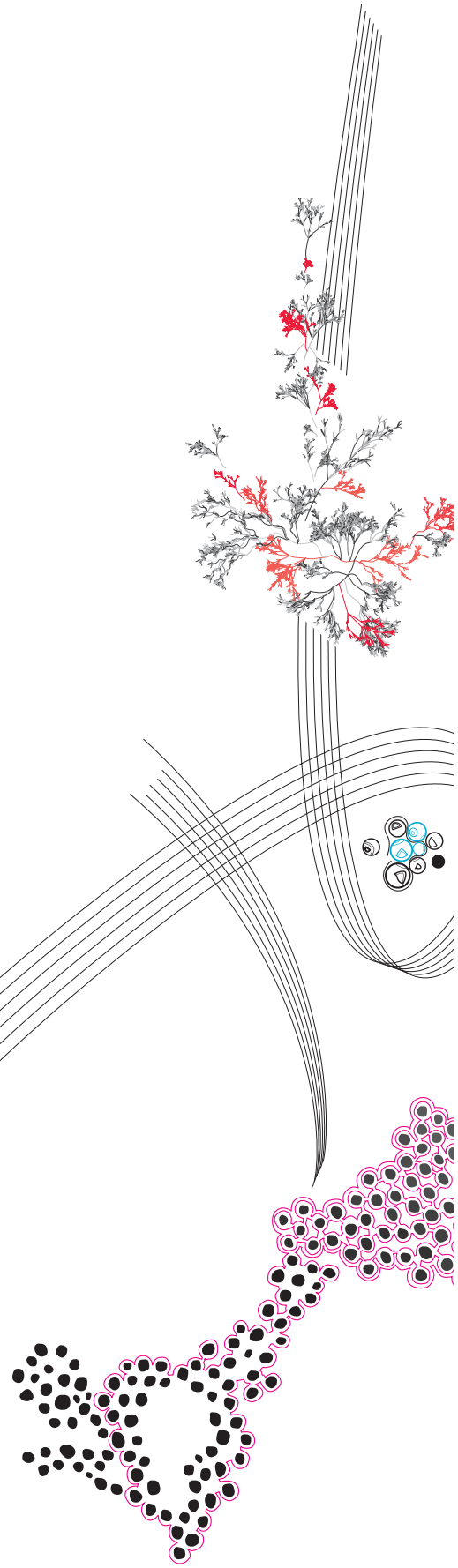

MSc Business Information Technology

The FAIR-iT framework: Towards the adoption of the FAIR principles in an enterprise context

Marjolein Kerssens

Supervisors:
Robin Effing - Faculty of Behavioural, Management, and Social Sciences
Luiz Bonino da Silva Santos - Faculty of Electrical Engineering, Mathematics, and Computer Science

August 12, 2023

MSc Business Information Technology
Faculty of Electrical Engineering,
Mathematics and Computer Science,
University of Twente

UNIVERSITY OF TWENTE.

Preface

After six years, my student life has come to an end. I have thoroughly enjoyed my time at the University of Twente. Starting out with the Creative Technology bachelor, where I quickly discovered my interest in business IT improvement and innovation. I am thankful for my academic and professional development and progress, as well as my personal development. The thesis that is presented is the result of six months of hard work and I am happy to say that I am quite satisfied with the result. There are many people who have contributed to the thesis in the past few months and I would like to thank some of them in the following:

I would like to start by thanking my university supervisors Robin Effing & Luiz Bonino da Silva Santos. Your feedback has been of crucial importance in my achievement of finishing the thesis. Both of you have provided interesting insights and have pushed me to do my best. You have always supported my ideas and were open to discussion, but also provided autonomy, which I greatly appreciated.

Secondly, I would like to thank all of my company colleagues who took the time to provide feedback or discuss my worries about the thesis. A very special thanks to Claudia, who was my company supervisor. Thank you for always making time in your schedule to address my worries and for giving me access to your network, which has greatly aided me in finding participants for my research. Additionally, I would like to thank my fellow interns with whom I could always have a laugh even if there was very little to laugh about. They kept me motivated with frequent coffee breaks where many kind words of encouragement were exchanged.

Last, but definitely not least, I would like to thank my friends and family, especially my partner and parents, who were always there to guide me when I needed it most. Your words of encouragement have enabled me to push my limits and produce this thesis. It would have been impossible without you.

Abstract

The FAIR guiding principles are findable, accessible, interoperable, and reusable and they can be used by organisations to govern their data assets through human- and machine-actionability (M. Wilkinson et al., 2016; GO-FAIR, 2018). The performed exploratory and structured literature reviews suggest that the FAIR guiding principles are mostly used in research organisations, but research about their use in an IT enterprise context remains absent. The benefits for research organisations have been written about abundantly, but their benefits to the IT consultancy sector remain uninvestigated, at least to the best of the author's knowledge.

The aim of the research at hand is to investigate the existing gap in the use of the FAIR guiding principles in IT consultancy storage and archiving of engagement data. A structured literature review was performed and 4 case studies were conducted. The results from the exploratory case studies show that there is room for the implementation of FAIR, as the employees at the selected enterprise experience challenges in the process of storing and archiving engagement data and finding data back once it is archived. As a result of the systematic review and case studies, the FAIR-iT framework was developed with the following principles in mind: findable, accessible, interoperable, reusable, internal, and trustable. Importantly, the last two principles have been added to emphasise the benefits that companies can gain from implementing the FAIR principles according to their corporate standard and are more focused on the governance of data. The added principles are not separate from the foundational four principles, but are supportive and only emphasise what the original principles already encompassed. The framework also provides a data preparation flow to help employees feel empowered in their tasks and broaden their skills so that they can take on the responsibility of caring for engagement data. Finally, the framework also provides a plan for change management once an enterprise decides to implement the FAIR-iT principles in its workflow as literature has shown that user acceptance is of importance in implementation projects (Akrong et al., 2022; Shah et al., 2011).

The findings of this study implicate that enterprises can benefit from the adoption of the FAIR principles, but that there is a different angle to be taken than when they are implemented in research organisations. Enterprises are more reticent about their processes and data to maintain a competitive advantage, and because of contractual agreements with relation to the data of their clients. This is why the FAIR principles will be useful for internal use at first. However, as several case study participants noted: there might be a future where there is room for collaboration rather than competition, and that is where the use of the FAIR-iT principles will flourish; once they become the FAIR-T principles.

Keywords: FAIR guiding principles, FAIR-iT framework, engagement data storage, metadata, enterprise data, change management

Contents

List of Figures

List of Tables

Acronyms

AI Artificial Intelligence

EOSC European Open Science Cloud

ERP Enterprise Resource Planning

GDPR General Data Protection Regulation

FAIR Findable Accessible Interoperable Reusable

ML Machine Learning

OD Open Data

PMO Project Management Ondersteuning (Project Management Support)

RDF Resource Description Framework

Chapter 1

Introduction & Thesis framework

This chapter first introduces the topic at hand and describes the problem statement. Additionally, this chapter contains sections where the research design, literature review method, and research method are explained.

1.1 Introduction

The FAIR guiding principles are four foundational principles that have the goal of enabling long-term care of digital assets so that they can be reused for future research (M. Wilkinson, Dumontier, & Aalbersberg, 2016). The principles are Findable, Accessible, Interoperable, and Reusable. Together, they should enable organisations to manage their data more effectively. Literature suggests that the current application for the FAIR principles is mostly on public and scientific data (GO-FAIR, 2018). The idea behind the FAIR principles stems back to the concept of Open Data (Murray-Rust, 2008). The concept of Open Science (which stems from the concept of OD) is supported by the European Open Science Cloud (EOSC) which has the goal of accelerating and supporting the transition to increasingly effective science. The goal of OD and the EOSC is to reuse data more efficiently and easily. This is beneficial to research organisations as it is likely to speed up the research process (Wise et al., 2019), but it could potentially be beneficial to enterprises as well, as data reuse is not something that occurs frequently but could benefit organisational processes (Labadie et al., 2020).

Enterprises gain an increasing amount of data on a daily basis (Wixom & Ross, 2017). In order to make these amounts of data advantageous, the data will have to be analysed and stored in a way that makes it possible to find it back, while ensuring that the data is kept safe from those without access (George et al., 2014). As the quantities of data are ever increasing, the repositories and archives from enterprises are increasing too to keep up with the larger amount of data. It is reported that between 60% and 70% of enterprise data is unused (Gualtieri, 2016). The kept data is often not used to its full potential as it is not well maintained, trustable, or of high quality. As more enterprises rely on data for their business practices, an increasing amount of employees have to become data citizens (or: someone who uses data on a regular basis). The requirements to encourage employees to increase their data usage can be found in the FAIR principles (Labadie et al., 2020). Better management of data is said to lead to a competitive advantage if the information coming from

the data is leveraged correctly (Russom, 2017).

Academic research on the implementation of the FAIR principles has been abundantly presented. However, much of the research is concerned only with their implication for research organisations (Wise et al., 2019; Katz et al., 2021; Nicholson et al., 2023). As far as the author's knowledge goes, there is little research on the use of the FAIR principles in an IT consultancy and audit context. This research gap, combined with the ever-growing amount of data that enterprises have at their disposal, creates an opportunity for research. The FAIR principles could be advantageous to enterprises as they provide guidelines on how data could be stored in a way that makes it findable, accessible, interoperable, and reusable (Labadie et al., 2020), which will be important as company archives have to store more data than ever before. The research at hand aims to investigate whether the FAIR guiding principles can benefit enterprise organisations in the same way as the principles benefit research organisations with the goal of unifying and streamlining the enterprise data storage and archiving process through the use of the FAIR principles with as little employee resistance as possible by designing an artefact that satisfies these requirements. The scope of the thesis will be the storage and archiving of project data in the form of engagement files for consultancy and audit departments, as the specific enterprise the thesis was written at identified this to be a challenge.

1.1.1 Thesis structure

The thesis utilizes the following structure. Section 1.2 describes the research design, the steps that will be undertaken, and the framework used. This is where the research questions can be found. Also, it explains the research and data gathering in detail. Chapter 2 provides extensive background knowledge and summarises academic sources available in the context of FAIR principles. 3 shows the results from the interviews and the draft version of the framework, which will later be validated. Chapter 4 describes how the validation of the draft of the framework was performed and what the results were. Chapter 5 contains the final version of the framework and the accompanying advice for change management within the organisation the thesis was performed at. Chapter 6 discusses the implications of the results for academics and practice. Here, limitations will be discussed as well. Chapter 7 concludes the thesis and provides key contributions.

Please note that because of a non-disclosure agreement, the data used from the company will not be shared as part of the thesis. This includes enterprise documentation and names from persons, departments, or the enterprise itself. The necessary information has been made anonymous.

1.2 Research Design

This thesis contains descriptive research and design research, which will be of an exploratory nature. The descriptive research is conducted by doing a systematic literature review and case study which address the knowledge questions that are critical for understanding the state-of-the-art of this field and aid in defining a framework for the preparation of the adoption of the FAIR principles. During the thesis, the

design science framework described by Wieringa (2014) will be followed. This framework was chosen as it provides guidelines on how to conduct design research but also allows for the answering of knowledge questions in support. Wieringa (2014) describes the goal of a design project to be the (re)design of an artefact in order to improve its contribution to the achievement of a larger goal. The proposed design cycle describes the process of a design research project which comprises the following phases: Problem investigation, treatment design, and treatment validation. This then flows into the treatment implementation and implementation evaluation, as the design cycle contributes to a larger engineering cycle. The latter two phases will not be part of this thesis as the goal is not to implement a framework, but to design it, and one cannot evaluate a framework that has not been implemented. It is important to note that this cycle is iterative and it could be the case that the steps are taken several times in order to ensure that the designed artefact satisfies the requirements. The design cycle can be found in figure 1.1.

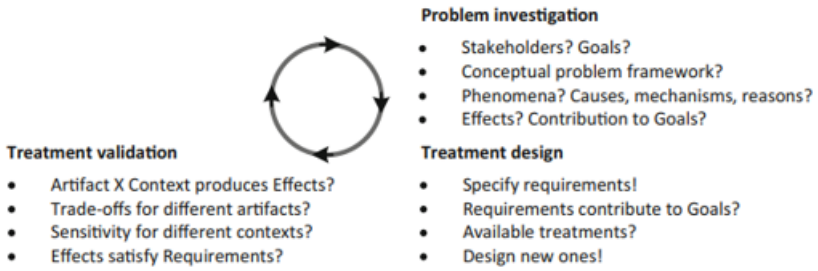

Figure 1.1: Design cycle adapted from Wieringa (p. 28)

Wieringa (2014) mentions that his design cycle follows steps that are similar to the design science research methodology that Peffers, Tuunanen, Rothenberger, & Chatterjee (2008) designed. The choice was made to use the methodology by Wieringa (2014) rather than the one from Peffers, Tuunanen, Rothenberger, & Chatterjee (2008) as the former encompasses a conceptual framework that also includes the design of research problems and accompanying questions, together with research methods that can be applied throughout the cycle.

1.2.1 Research objective and questions

Wieringa (2014) proposed a template for the definition of design problems (which he also refers to as technical research problems). Following the proposed template, the objective of this thesis can be defined as *to improve <the method IT consultancy and audit departments use for data storage> by <developing a framework> that satisfies <the FAIR data principles and highlights the change management process> in order to <guide the preparation of going FAIR and thereby improving the process of finding and working with data from past projects>*. The template can be seen in figure 1.2

Improve <a problem context>
by <(re)designing an artifact>
that satisfies <some requirements>
in order to <help stakeholders achieve some goals>

Figure 1.2: Wieringa's template for technical research questions (p.16)

From this design research objective, the following main research question has been formulated: *How can the applicability of the FAIR data principles in an enterprise context be enhanced, specifically so that a framework can be developed that allows organisations to prepare the adoption of FAIR into the data storage process of engagement files in IT consultancy and audit departments?*

Since the goal is to make adaptations that allow for the implementation of the FAIR data principles within consultancy and audit departments, several descriptive knowledge questions must be answered. The choice for these consultancy and audit departments was made because these people are often already used to working with data, so they already are data citizens. Next to this, a fitting implementation framework must be developed using a technical research question. The choice for a framework over a procedure is because a framework is less specific and rigid. The goal is to create a set of approaches that will contribute to the implementation of FAIR, not to have a list of steps that must be followed. The first knowledge question is defined as:

1. What challenges should the implementation of FAIR solve and how is this currently orchestrated in an enterprise context?

Firstly, the state-of-the-art use of the FAIR data principles in a business context should be researched. The reason for this is that exploratory literature research has shown that the principles are mostly used by universities and other research organisations, but not yet by businesses. For this sub-question, it is important to note that this will likely not be only consultancy companies but also other sectors. The state-of-the-art will be reviewed in the literature review, but there will be no separate research question dedicated to the topic as it should not be the main focus of the research.

The second question aims at identifying practices that can be deployed to effectively manage data using the FAIR data principles. This will include existing strategies and frameworks, as well as best practices.

2. What frameworks and tools have been developed to aid the adoption and implementation of the FAIR principles and thereby help organisations go FAIR?

The main research question highlights the desire for an implementable framework. Answering this research question will aid in identifying actions and risks that consultancy companies might encounter when making the step towards FAIR while

obstructing the implementation process as little as possible. This is why it is important to not only consider the benefits a department can reap from implementing the framework but also assess the impact it could have on current business practices.

The last question focuses on the change management aspect of the preparation for implementation. The final research question is as follows:

3. What change management steps should enterprises undertake to guide the adoption and implementation of the FAIR principles?

The implementation of a new framework will require a change management plan as this is crucial to create an implementation with as little resistance as possible. The final sub-question focuses on this aspect. It is well-known that guided change management is a critical success factor in any business change that affects employees on a large scale. Therefore, managing change is crucial to the success of the framework that is to be developed (Kerssens, 2023) [1].

1.2.2 Research Model

The combination between the research questions and the research design is presented in the figure below. Data will be gathered in several ways during the research. The first way is a systematic literature review. With this method, some of the sub-questions will be answered. Then, a case study will be conducted in order to investigate the requirements to ensure that the framework is fitting to the style of working within the IT consultancy firm the internship is performed at. The case study will be held with managers and senior managers, as they have more insight into the process level compared to other employees. The combination of these two methods will create the draft for the framework, change management plan, and accompanying road map. The reason why it is not only based on literature is the lack of empirical research in relation to this topic. The outcome of the systematic review and the case study will seek to answer the knowledge questions and gain a clear perspective on the problem context. This will then ultimately draft the first version of the framework.

The third part of the cycle is to validate the created artefact. This means that the draft of the framework will be validated. This will be achieved through a series of interviews with different participants than the case study to see whether the framework fulfils the expectations and needs of the people who will work with the framework the most. First, participants are asked to review the draft to evaluate the benefits of the framework for their department or project teams. Secondly, they will be asked whether they think the framework is something they can work with once it is implemented and whether the change management plan fits with their expectations. The choice to do this was made because research has highlighted that employee involvement is a deciding factor in change management, as was mentioned previously. The input from the target group will then alter the draft framework in order to create the final version of the framework. In figure 1.3, the vertical arrows at the bottom indicate the phases of the design cycle from Wieringa (2014), whereas the boxes above indicate what actions are performed during that cycle.

[1] This statement is the result of previous research performed in an earlier issued report by the same author

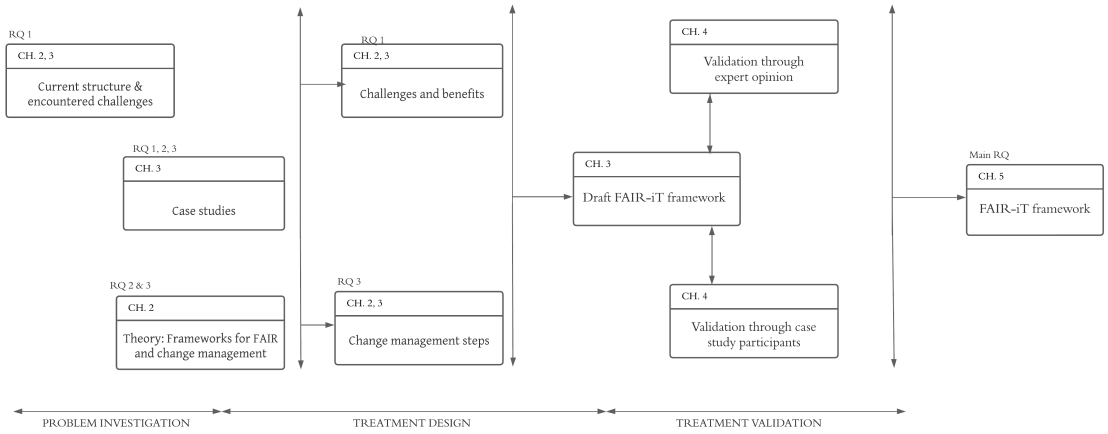

Figure 1.3: The proposed research model

The use of the design cycle in this research as proposed by Wieringa can be seen in more detail in figure 1.1.

Table 1.1: Applied design cycle phases

DESIGN PHASE	RESEARCH FOCUS	METHOD	CHAPTER
Problem investigation	State-of-the-art of FAIR tooling and application in different sectors	Literature review	2
	RQ1: What challenges should the implementation of FAIR solve and how is this currently orchestrated in an enterprise context?	Literature review, case study	2, 3
Treatment design	RQ2: What frameworks and tools have been developed to aid the adoption and implementation of the FAIR principles and thereby help organisations go FAIR?	Literature review, case study	2, 3
	RQ3: What change management steps should enterprises undertake to guide the adoption and implementation of the FAIR principles?	Literature review, case study	2, 3
Validation	Change management effectiveness	Expert opinion	4
	Requirements	Expert opinion	4

1.3 Literature review method

This chapter contains a literature review of concepts that relate to the use of the FAIR principles in a business context. This will be split up into two parts. One will describe the use of the FAIR principles in other industries than business. The second part will present existing frameworks related to the use of the FAIR principles. Lastly, a conclusion from these two parts will be drawn, which will serve as one of the base steps for the draft of the to-be created method.

To ensure that relevant literature concerning the FAIR principles is selected, a structured literature review was performed using the strategy proposed by Kitchenham (2004). The process will have an iterative nature, where the inclusion and

exclusion criteria will be adapted based on the search results. Please take note that the search for the systematic literature review was carried out in week 22 of 2023 and may differ if the search would be performed again. The search protocol used can be found in Appendix A. The search protocol is a detailed description of the used search terms and the criteria for inclusion and exclusion. This is done to make sure that as little bias and as much coherency as possible exists when selecting relevant literature. The search terms were initially created through an exploratory literature review, but were later adapted in an iterative way to ensure that all relevant literature was captured. An example of this is the exclusion criteria of the threshold of 15 citations. The threshold was moved up and down in several databases to see what it would do to the results that showed up. The reason for this is the suggestion that some of the relevant literature might be too young to be cited as often as older papers, whilst being of relevance to the topic at hand.

The papers were selected based on scanning of the titles of all of the results of the selection that was made after applying the inclusion and exclusion criteria. All papers that did not seem to have a relevant title were excluded immediately. Then, the abstracts and key words of the left over papers were read, and once again all papers that had abstracts that seemed irrelevant were removed from the selection. As a last step, the selection of papers were read through fully, and based on the read-through a final selection was made. From the final selection, sources were checked and some sources went through an iteration of the method to see if they would fit in the review. The papers that were part of this final selection are the papers that were used in the literature review as all of the contents were relevant to the topic. The selection steps were carried out with a conservative nature, meaning that when doubt existed about the relevance of the paper, it would be taken to the next step to ensure that no relevant literature would be missed. This resulted in a total selection of 15 papers The process is visualised in figure 1.4.

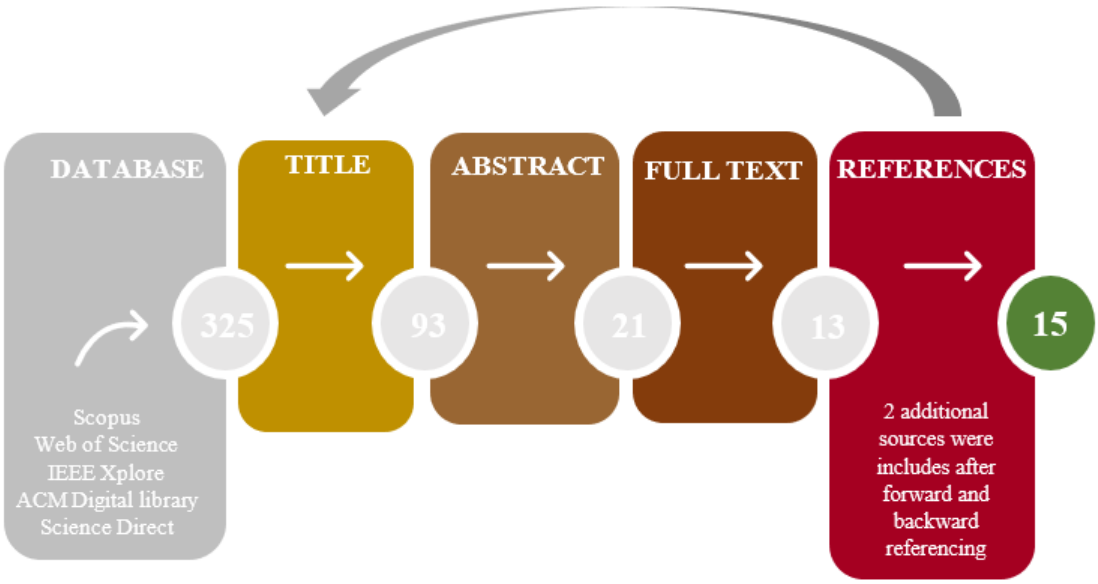

Figure 1.4: Structured literature review method and output

The quality of the literature is an important factor in systematic reviews. This is why only certain academic databases were researched and the selection was made from conference and journal papers only. However, it is important to note that the inclusion of literature that is not peer-reviewed (yet) can be important in research that is related to software engineering, according to Garousi, Felderer, & Mäntylä (2017). This is especially important when there is not a substantial amount of literature from the field or research, which is the case with the FAIR principles in relation to business. Grey literature (literature that is not peer-reviewed) might provide state-of-the-art ideas and concepts that might be too new to be reviewed already, which may help avoid certain types of bias like publication bias. There was little academic literature available that focused specifically on the use of the FAIR principles in business sectors. The use of grey literature in combination with academic literature makes this literature review a "multi-vocal" literature review. A risk of conducting multi-vocal literature reviews is that the quality of the papers might be lower than standard. Therefore, only first-tier grey literature was used that had high credibility, examples include grey literature found in books and whitepapers.

During the search as mentioned in Appendix A, only literature from 2016 onward was selected as the FAIR principles did not exist before that. Later, the possibility arose to use older research as a result of the reference research as those papers presented a deeper understanding of the background of the newer research. However, all of the selected reference papers fell into the chosen time frame. In week 26 of 2023, another exploratory search was performed using Google Scholar, to ensure that no new data was published during the writing of the thesis. This did not yield any new results, so the list of papers is considered complete. The list can be found in Appendix B.

A total of sixteen papers were selected for the final list. Of these papers, there are six journal papers and seven conference papers. The list was completed with one article and two white papers. Interestingly, most of the papers are not based on empirical research. Instead, many of the papers are based on expert opinions, which are often personal opinions from the authors themselves. A small selection of the papers was purely based on literature. Almost none of the papers are based on empirical evidence such as interview results, case studies, or a survey. This is likely because research in this sector has little opportunity to work with empirical evidence, and there is often a significant reliance on expert opinion, according to Kitchenham (2004). However, some researchers performed tests with the use of scripting and simulation.

The development of the framework consists of several parts, some of which include: a problem investigation, research about usable frameworks, methods for implementation and management, and lastly the change management aspect. The results of the literature review will be split up into these important aspects. The first three parts will stem from the structured literature review, whereas the last part will be taken from a previous report by the same author as well as some state-of-the-art information on change management.

1.4 Research Method

1.4.1 Case study

The exploratory literature review served as the basis of the theoretical background provided in chapter 2. During the exploratory literature review, it has become clear that a lack of empirical research in the field of FAIR in enterprises exists. Therefore, the gathering of empirical evidence is imperative in the context of this research. This should allow the author to base the to-be-created method in a real-world context, thereby increasing the reliability of the conclusions. The choice was made to investigate the requirements of a framework further by conducting a case study. Additionally, the second goal of the case study is to investigate the requirements of a change management plan, as well as the general opinion about change within the organisation. The case study follows a replication logic while using a multiple-case holistic design, as proposed by Yin (2018). The FAIR principles are the global unit of analysis per enterprise and only one enterprise was analysed with employees from various departments to see if their answers will be comparable. Evidence was collected in three ways:

- *Semi-structured interviews:* The main input for the data collection was done through a case study. A case study was conducted with employees from different departments related to IT consultancy and audit within the same enterprise. Case study participants were selected based on position, years of experience, and project experience related to the topic. For example, the years of experience the participants have ranges between X and Y. This choice was made because a spread in the amount of experience is considered desirable to provide insights from employees with a lot of experience as well as those who do not have a lot of experience yet. The choice for choosing participants from several positions is to see if the challenges they encounter are spread evenly across the firm. The case study was held in individual sessions of about an hour through Microsoft Teams. The first part of the case study was about problem investigation and their view on a solution, after which the FAIR principles were introduced. Here, participants were shown results from the literature research as well as a short explanation of the FAIR principles. This choice not to show the FAIR principles immediately was made so that the interviewees were not instantly steered in the direction of FAIR, as their views on other possible solutions are important to the study as well. Lastly, the respondents were asked about their view on change management within the organisation and their opinion about what matters when implementing change.

- *Additional documentation:* Additional documentation was obtained in the form of presentations, documents, and figures of what the enterprise-developed change management strategy looks like. This information is considered confidential and will only be described indirectly so as to not violate the regulations with respect to sharing of this information.

- *Exploratory conversations:* Before the process of official data collection started, the author searched for a specific scope for the thesis to take. This was done

through informal conversations during exploratory meetings. Additional remarks that were interesting for the empirical data collection were written down as long as there was no mention of confidential information. The employees selected for these conversations all worked intensively with the data collection process and identified themselves as knowledgeable on the topic through their job titles and description available on the enterprise intranet. A total of six exploratory conversations were held in weeks 17, 18, and 19 of the year 2023.

Four experts were selected with the goal of achieving saturation in the data set. A data set is said to be saturated once no new data that provides additional information is found (Glaser & Strauss, 1967), hence when the participants do not provide new challenges or solutions in their answers. In the case these four participants would not have been enough to rule out new answers, more participants would have been added to the pool so that saturation would have been reached.

Analysis of case study results

The sessions were recorded and transcribed using Microsoft Teams. All participants gave permission for this. After the interviews were completed, the recordings and transcripts were cross-checked to see if everything was correctly transcribed. Changes were made in the cases where the transcript was not correct. In some instances, the in-between conversation did not contribute to answering the case study questions (an example of this is the question if participants could see the shared screen). These pieces of the transcript were removed for clarity, the rest of the transcript is precise. Some parts of the transcripts have been removed as they contained information about clients or the enterprise itself. Due to the non-disclosure agreement, these details have to remain anonymous. The average length of a transcript is eleven pages. The transcripts were not sent to the participants as most of them did have time to check for feedback. The interviews all lasted between 45 and 55 minutes.

The case study was analysed using ATLAS ti, which is a qualitative data analysis software. The concept of open coding was followed (Bryman, 2012). Relevant sentences and paragraphs were assigned certain codes that related to key topics of the research. During and after the coding process, certain codes were added, merged or deleted. Sentences and paragraphs were assigned one or multiple codes based on the contents. The relationships between the code categories were taken from the relationships between codes in their subgroups. An overview of the coding can be found in Appendix D. Information from additional documents was not coded as it already related to very specific codes.

1.4.2 Validation

In order to validate the created artefact (the framework, in this case), the artefact should be validated. Venable, Pries-Heje, & Baskerville (2016) describe different strategies for the validation of design science research. The first distinction is made on the basis of the goal of the validation. In formative evaluation, the goal is to improve the artefact based on the validation results. In summative evaluations, the goal is to test whether the artefact matches expectations. The methods used for the evaluation are the basis for the second distinction. Naturalistic evaluations intend

to validate the artefact in the real environment where it will be deployed. Artificial evaluations, on the other hand, intend to evaluate the artefact using laboratory experiments and simulations. The latter is not applicable to this thesis and will not be considered as a validation strategy.

Next to the categories of evaluations, Venable, Pries-Heje, & Baskerville (2016) also define a variety of strategies for evaluation. The strategy that has been selected for this research is the *"Human Risk and Effectiveness"* strategy. This is chosen because it has a focus on a design risk that is related to social difficulties or users, there is a large possibility to evaluate with real users in the real context, and lastly, because the goal of the evaluation is to establish the benefit of the artefact in real situations. The chosen strategy emphasises formative evaluations early. Likely starting out with formative, artificial evaluations, but moving towards formative naturalistic evaluations quickly. Near the end of the development of the artefact, and thus the strategy, the focus will be on the evaluation of the effectiveness of the artefact, thereby moving towards naturalistic summative evaluations. As the goal of this research is to ascertain whether the FAIR principles can benefit enterprises, the focus of the research steers away from the technical aspects and more into the organisational and social domain. This is why the chosen strategy is likely to fit well with the goal of the thesis.

Expert opinions

The framework for FAIR implementation was validated through expert opinions from experts in the field of IT consultancy who were also working at the case company. This means that the validation was performed internally. According to Wieringa (2014) it is important to select experts who would work with the artefact intensively and have extensive knowledge of how the artefact will interact with the problem context at hand. They must predict what effects will arise. The experts were selected based on years of experience and department. The goal was to have a range of years of experience to prevent only interviewing employees who had just started and might not have encountered many challenges yet. Only employees from consultancy and audit departments were selected for the validation as these departments are the ones the framework should be applied to. None of the experts were involved in the research before so there would be an unbiased view on the framework.

All interviews were recorded through Microsoft Teams and then transcribed. The interviews had a semi-structured approach and were similar in order to the case study. First, the principles were presented in the form of an overview, after which the data processing flow was shown, followed by the change management plan. As the goal of the interviews is to validate the findings and see if aspects of the framework or change management plan had to change, no coding was applied to the transcripts. The process of validation was not of an iterative nature. All of the experts were presented with the same information about the framework and change management plan. Then, after the interviews, the framework and plan were adapted according to the reactions and comments from the experts. The choice not to make the process iterative, so the framework changing in between every expert interview, was made because that could lead to a situation where the framework and change management plan would be continuously adapted back to previous versions which

would lead to difficulties regarding time constraints.

Previous respondents

The participants from case study were requested to take a look at the draft framework. The draft framework was shared with them through email. They were asked to provide written feedback on the framework to see whether the requirements they expressed were met. Unfortunately, only one expert was available at the time of validation.

Chapter 2

Theoretical background

This chapter consists of the theoretical background of the thesis. First, background information about the various topics will be given. After that, a literature review of state-of-the-art concepts, FAIR tools, and literature on change management will be provided.

2.1 Background

This chapter presents information about key concepts related to the thesis. These will be separated in the question of what the FAIR principles actually are, their benefits and challenges, and the concepts of open science and data stewardship, which are both closely related to the FAIR guiding principles.

2.1.1 What are the FAIR principles?

The "FAIR Guiding Principles for scientific data management and Stewardship" were published in 2016. According to the GO FAIR website, the goal of the FAIR principles is to provide guidelines to improve several aspects of digital assets. The foundational principles are Findability, Accessibility, Interoperability, and Reusability. If all aspects are implemented correctly, the data should be easy to work with for humans and Artificial Intelligence (AI) agents. The four principles can be split up into fifteen guiding principles, which will be described below. Additionally, they can be found in figure 2.1. The principles refer to three categories of entities: data, metadata and the infrastructure of the data. This ensures that data is well described in itself and a meta-variant, while also ensuring technical compliance, such as a location in a system (GO-FAIR, 2018). Metadata has been described as "data that describes other data", which is rather vague. Metadata contains no information about the contents, but rather on how it was gathered, when it was created and the size of the data set. This should enable users to easily find, manage, and use the data set (Farrier, 2022). The FAIR principles advocate that metadata schemes require standardisation in order to simplify the process among sources, which is something that influences research communities as they might have to change the way they structure data. Adding to this, the GDPR requires that metadata is verified, and whether metadata provides information on the data subject and processing policy (Landi et al., 2020). FAIR handles this requirement by treating all pairs of data and metadata in isolation. The metadata is considered to be the descriptor, and the data itself is always the thing it describes within the context. This leads to

a situation where even if the context changes, the link between the data and the metadata does not, even if data suddenly becomes metadata in the renewed context (Jacobsen, de Miranda Azevedo, et al., 2020). Extensive metadata has also been identified to improve the comparison and bench-marking of certain approaches, such as Machine Learning. This is because it can provide a baseline that can be used to group certain requirements with data sets or algorithms. This could greatly benefit research as it could speed up the process of data collection, which is something that FAIR advocates as well (Katz et al., 2021).

Figure 2.1: The FAIR foundational principles and their guiding principles

Findability

An important step in working with data is finding the data that is needed. GO FAIR describes that both data and metadata should be easy to find for all entities that have authorisation. Metadata which is machine readable is crucial for the discovery of data sets by AI and Machine Learning (ML) agents. The Findability principle can be split up into multiple counterparts that, together, make a data set findable. The action points for this principle are focused on the identification of a data set, the use and proper description of metadata, and the indexation of (meta)data in a resource that can be searched (GO-FAIR, 2018). Principle F1 mentions that there should be a "globally unique and persistent identifier". This means that the identifier should be unambiguous across all data sets, and not just local ones, all while never being reused in a different context and ensuring that it always is linked to the same resource, which is the meaning of persistence (Jacobsen, de Miranda Azevedo, et al., 2020). This can be done through the use of a Resource Description Framework, which is a globally-accepted framework for the representation of data so that it can be interpreted by machines globally (M. Wilkinson et al., 2016).

Accessibility

If the first principle, Findability, is implemented correctly then the next step is that a user can access the found data. This process could depend on authentication

and authorisation, depending on the source and content of the data. With this, it is important that (meta)data is retrievable through an identifier and that the communications protocol that is used is open, free, and implementable on a universal scale. Next to this, it is of significant importance that the metadata of the data set is still available, even if the data set itself has been deleted (GO-FAIR, 2018).

Interoperability

It often occurs that data needs to be integrated with other data sets. Next to this, data must be added to a system or application. In this case, the data must be interoperable with the standards of said system or application, while also being interoperable with several workflows for storage and processing. This must, however, not interrupt the principle of accessibility, where it is important that integrity is ensured and unauthorised access is prevented (GO-FAIR, 2018). This foundational principle is especially important as users spend enormous amounts of time making sense of the data sets they need and finding accurate ways to make them interoperable. This is often the result of ambiguous descriptors that are occasionally not machine-interpretable, thereby making it impossible to automate the task (Jacobsen, de Miranda Azevedo, et al., 2020).

Reusability

This is described as the most important goal of the FAIR principles, to ensure data reusability. This can be achieved through thoroughly-described data and metadata. This ensures replication and the combination of data sets. Data is reusable if it has accurate attributes attached to it. This can be done through licensing, provenance, and community standards (GO-FAIR, 2018). Principle R1 seems very similar to principle F2 as both are related to the richness of (meta)data. They are, however, not the same. The focus of F1 is related to the desire to enable effective search and query, thereby making it findable. R2 focuses on the need for machines and humans to easily assess if a source is suitable for reuse (Jacobsen, de Miranda Azevedo, et al., 2020).

Everything FAIR is not

Identified by Mons et al. (2017) is the notion that some aspects of the FAIR principles can be ambiguous in their interpretation. They made a list of everything that FAIR is not. Below, a summarised version of the list can be found for clarity on what is, and what is not included in the FAIR principles.

- The FAIR principles are not a standard. They are guidelines and should be seen as such, as they are open enough for interpretation to allow several approaches to implementation.
- Although linked to, they are not the same as a Resource Description Framework (RDF), the semantic web, or linked data. Some desires might overlap, but it would be wrong to categorize all of these entities under the same name. However, it is important to note that the FAIR principles can be used in collaboration with all of the other concepts.

- Some of the principles that fall under the guiding principle of "accessible" might cause confusion with potential users. Something that is accessible is not the same as it being open. The notion that a data set is accessible does not necessarily mean that everyone can have access to it. Certain accessibility and authorization rules will apply. Certain parts of datasets can be shielded, especially if that data concerns personal details. However, the license must be clear and machine-readable.

2.1.2 Benefits of the use of FAIR

Research has shown that the implementation of the FAIR guiding principles can have significant benefits for research institutions and employees. Russell (2021) highlights the ability of researchers to verify their research findings through the publishing of data and models so that other researchers can recreate the research and its findings. Next to this, it is mentioned that through the use of FAIR, institutions would be able to save a large amount of money on data collection, as there is a possibility to use already gathered data. Wise et al. (2019) agree with this point and state that time-to-value will show a significant decrease, all while productivity is likely to rise. Russell (2021) underlines the notion that this could be particularly relevant for publicly funded research. Lastly, it is argued that an increase in the availability of data should accelerate discoveries and enhance collaboration.

Wise et al. (2019) add other benefits to the list and points out that the use of the FAIR principles could lead to an increase of data running through the research value chain, thereby increasing semantic alignment and integration possibilities, all while increasing productivity in the research and development pipeline. Next to this, it is articulated that time-to-market for discoveries could be reduced and that it will become possible to develop more personalised medicine (in the case of biopharmaceutical research). Lastly, it is emphasized that the FAIR principles will enable data sharing across institutions and firms.

Nicholson, Kansa, Gupta, & Fernandez (2023) declare concrete advantages for archaeological research. These can be summarized as the creation of data summaries, specialist databases, and transactional databases. They state that not only researchers but also students will benefit greatly from the creation of the aforementioned databases, as less time is wasted on data transformations. Lastly, it is suggested that making data FAIR allows researchers to add to other's data sets, thereby enriching existing data with new samples.

All in all, the implementation of the FAIR guiding principles is expected to lead to increasingly rigorous data management and data stewardship of digital resources, something that large research communities could benefit from (M. Wilkinson et al., 2016).

2.1.3 Challenges in the use of FAIR

Landi et al. (2020) identify a crucial first challenge with the use of the FAIR guiding principles. The point of FAIR is that data is kept in repositories so that it can be accessed at a later point. In 2018, the General Data Protection Regulation

(GDPR) was deployed in Europe. From that point onward, companies and research institutions are required to protect the data of natural persons with regard to the processing of said data more than they were in the past. The "R" of FAIR stands for reusable, but this might be a challenging wish in combination with the GDPR as people might not (be able to) agree with other parties using their personal data, while this is the point of FAIR. This could also trouble the process of licensing, as data might not be accessible for a variety of reasons, such as the GDPR but also other forms of use agreements or the need to transform the data to ensure it is machine-readable. Nicholson, Kansa, Gupta, & Fernandez (2023) agree with this point and add the notion that user licenses should only be given out if this is deemed ethical, and that this process might be influenced majorly by regulatory agencies.

The implementation of FAIR will mean that firms have to adapt their processes and workflows, thereby incurring costs. In order to make the implementation successful, executive management needs to be convinced that this investment is urgent and worth it. Next to ensuring that management is on board, it is important that company culture changes along with the pipeline. The gatekeeping of data must be eliminated before the FAIR guiding principles can be beneficial. Enabling cultural change is often difficult, as a lot of it depends on how the teams operate internally and among each other. Involvement of stakeholders, such as employees, is of significant importance to lower the challenge of implementation (Wise et al., 2019).

M. Wilkinson, Dumontier, & Aalbersberg (2016) highlight a more technical challenge. Repositories often accept a large variety of data types in different formats, but there is no attempt to integrate the different data sources, all while placing little restrictions on the descriptors of said data. Because of this, the data ecosystem in the repositories seems to be straying away from centralisation, while that is exactly what is needed. The data and repositories are becoming more diverse and less integrated, thereby creating problems with reusability and interoperability.

Jacobsen, de Miranda Azevedo, et al. (2020) mention challenges for almost all principles, thereby giving a clear overview. Starting with the "F", it is identified that it often occurs that identifiers for data sets do not survive the termination of a project. This leads to data being difficult to find, as there is no longevous identifier. The next challenge occurs with the use of rich metadata, as there is no measure for when metadata is rich enough, there is no so-called minimum "richness". Adding to this, it is noted that there is a variety of metadata models that are all machine-readable, but not necessarily interoperable. Lastly, it is challenging to define a single source for indexing of metadata, and there is no singular way to execute a search. These two combined lead to the challenge of searching across a broad spectrum, as most search engines need specific pointers. Moving on the the "A", Jacobsen, de Miranda Azevedo, et al. (2020) mention similar arguments as Landi et al. (2020). Access is a challenging topic and many protocols are not as clear as they should be. Adding to this, it is described that all communities are to describe a persistence policy for their metadata, even if the data itself is not available anymore. This could lead to decreased stewardship. For the "I", the only challenge that was identified is that terminologies might not be the same, which makes interoperability hard. Machines in certain regions and communities might

not have the authority to access definitions, or data sets as a whole as described by Shanahan & Bezuidenhout (2022), to which they refer as "geoblocking". Lastly, the "R" poses challenges in the field of licensing. There is a lack of well-defined relationships that can distinguish several types of licenses, such as licenses for the metadata versus licenses for the data itself. Lastly, Jacobsen, de Miranda Azevedo, et al. (2020) call for a global standard instead of community standards, to improve reusability and interoperability. Unfortunately, this is currently not the case and this poses problems in the aforementioned principles. They conclude with the argument that the interpretation of the guiding principles is not straightforward. They allege that the success of FAIR depends on whether data stewards want it to succeed.

2.1.4 Data stewardship

Data stewardship is a term that is closely related to the FAIR principles, as their full name is "The FAIR Guiding Principles for Scientific Data Management and Stewardship". Data stewardship is an approach to managing data, especially the data of individuals. The goal, therefore, should be to have a selection of management procedures that cover obtaining, storing, and analysing data, all while ensuring that the data is anonymized so that the identities of the clients remain anonymous. Additionally, there have to be methods and regulations for the use of the data. The data steward, therefore, has the task of protecting the data, managing the procedures and methods, and, thereby, acting as a guardian to ensure trust (Rosenbaum, 2010).

Plotkin (2020) denotes a different definition and goal of data stewardship. Data stewardship can be viewed as the operational side of data governance. It has as a goal to formalize accountability through the management of information resources so that employees can benefit from it with the best interest of the organisation in mind. An important note is that the data steward is a person or team that manages data on behalf of others, often a business. Data stewardship can help organisations improve their metadata through collection and documentation. Additionally, having data stewards could lead to improved decision-making. Overall, having stewards will lead to improvements in the quality of data assets and enables enterprises to use their data as a competitive advantage. The difference between the notation given by Rosenbaum (2010) and Plotkin (2020) is likely the perspective that is taken in the definition. Rosenbaum (2010) views the goals of data stewardship from the side of the individual, whereas Plotkin (2020) takes the perspective of an organisation.

O'hara (2019) describes that the use of a "data trust" can be beneficial when attempting to achieve data stewardship among organisations and their clients. They remark that a data trust is compliant with the law to ensure that data processing is performed in a trustworthy fashion. It is stated that a trust deficit exists between organisations and their clients, despite efforts such as the GDPR. It will be important for all organisations that work with data (both enterprises and research institutions) to ensure that data is handled in a trustworthy manner that is compliant with the GDPR. This is where said organisations can greatly benefit from a data steward, who can aim to warrant trust.

2.1.5 Open Data & Open Science

The concept of FAIR dates back to the emergence of Open Data (OD). Open Data is a term that is used to define how scientific data can be published and reused without a permission barrier and for free. The concept of reuse is crucial to OD, but this is often blocked by publishers. Open Data sources are beneficial for data-mining practices and automatically aggregating data from multiple files. Licenses can be used to gain access to the data sources (Murray-Rust, 2008). Braunschweig, Eberius, Thiele, & Lehner (2012) mention that the advancement of Open Data is held back by legal, administrative, and technical requirements. They propose the Open Data repository to lower the technical hurdle of providing open data. The easy reuse of data will not be possible without open standards. Open standards can be defined by six principles, which are: end-user choice, availability, no discrimination, no royalty, and no predatory practices or subsets. The challenge is that several open standards are available and that it is hard to unify all organisations to one standard (Krechmer, 2005, 2008). Open Science is built on the concept of Open Data, as it provides a repository platform where shared data sources are provided.

The final concept that is related to FAIR, is the concept of Open Science. The concept of Open Science is supported by the European Open Science Cloud (EOSC) which has the goal of accelerating and supporting the transition to increasingly effective science. Additionally, the EOSC should enable research organisations access to systems, services and shared scientific data sources. So, the EOSC could serve as the platform on which FAIR data from research institutions can be published. It is remarked that the term "open" could possibly cause discussion, as not all data should be open due to privacy and confidentiality concerns. Additionally, it has been remarked that a change in culture is required before sharing of data can be optimized. Organisations and individuals must shift from seeing data as a powerful resource to seeing data as powerful when combined with data from other sources. Only then will the EOSC and FAIR principles flourish to the maximum extent (Commission et al., 2016; Burgelman et al., 2019). Budroni, Claude-Burgelman, & Schouppe (2019) highlight that the EOSC consists of six action parts which cover architectural, technical, regulatory, and governance-related action points. Additionally, it is mentioned that the FAIR guiding principles could prove to be beneficial to the data storage process within the EOSC, depending on the result of the working group assigned to researching this.

Burgelman et al. (2019) call for a new principle: Open Science by Design. As an example, they use a variety of outbreaks of diseases that have been stopped at a historical speed due to research organisations sharing their information, thereby speeding up the process of vaccine development. However, they do note the challenge of lack of recognition and citations, both of which are important in the research industry. They remark that changing the system of rewards and incentives will be a crucial step in creating a truly open science cloud. Additionally, they highlight the need for data to be FAIR before the system will truly be open and question the effectiveness of the cloud if the data in there is not usable or interoperable. To speed up the technical and cultural change, they believe that a new method should be introduced: Open Science by Design.

More and more governing bodies and research institutions adopt the FAIR principles as their baseline for data storage. This has led to an increasing number of resources being (or attempting to be) FAIR. In Europe, the EOSC is the platform for Open Science data, but other parts of the world have been adopting their versions of similar platforms. Examples of this are the NIH in the United States and the Dakar Declaration on Open Science in African nations. One could argue that the adoption of the FAIR principles will lead to increased adoption of Open Science practices, and vice versa (Mons et al., 2017).

The exploratory literature research that has been used to create the theoretical background has shown that the FAIR data principles are mostly used by researchers to share their data sets with other researchers from different institutions. This notion is taken from the fact that, currently, the FAIR principles are closely related to Open Science initiatives. An interesting question to answer would be whether it is possible to implement the principles in a business context, specifically IT audit and consultancy, and what implications that would have.

2.1.6 Related framework - the TRUST principles

The creators of the TRUST principles argue that the FAIR principles do not describe the need for trustworthy digital repositories well enough and lack sustainable governance. This is why they propose the TRUST principles for digital repositories. They provide five principles that all contribute to the effectiveness of the digital repository. The principles can be seen in figure 2.2. The TRUST principles provide a reminder for data stakeholders to develop and maintain their data infrastructure and continue their responsibility of the stewardship of the data sources they provide (Lin et al., 2020).

Principle	Guidance for repositories
Transparency	To be transparent about specific repository services and data holdings that are verifiable by publicly accessible evidence.
Responsibility	To be responsible for ensuring the authenticity and integrity of data holdings and for the reliability and persistence of its service.
User Focus	To ensure that the data management norms and expectations of target user communities are met.
Sustainability	To sustain services and preserve data holdings for the long-term.
Technology	To provide infrastructure and capabilities to support secure, persistent, and reliable services.

Figure 2.2: The TRUST principles for digital repositories

The choice to use the FAIR principles instead of the TRUST principles was made because the TRUST principles have less literature related to them. This will make it more challenging to perform an in-depth literature review, which is something that serves as an important component of the development of the framework.

2.2 Structured literature review

2.2.1 What challenge does FAIR solve?

According to Jacobsen, Kaliyaperumal, et al. (2020), the FAIR guiding principles aim to enhance analysis possibilities for both humans and machines. This could

solve the problem of data not being the same as information, and those numbers are increasing exponentially. Many enterprises sit on enormous amounts of data that are not being used to their full potential as their navigation is time-consuming (Roszkiewicz, 2010). The implementation of the FAIR principles could aid in this through the implementation of a metamodel that allows for easier tagging of datasets. Currently, many employees are not unlocking the maximum value of the data they are working with as it is often not well-maintained and occasionally does not have a trusted source (Labadie et al., 2020). Juty et al. (2020) add to this by asserting that large amounts of research data get lost within a small amount of time. FAIR should aid in this loss as it forces institutions to properly use identifiers that are unique to a certain data set, thereby preventing that data gets lost to a certain extent. Those unique identifiers could help humans and machines to find data sets more easily. Additionally, large amounts of data sets are often challenging to access and reuse because of the limited access or because they are difficult to find (Gryk et al., 2019).

Another challenge related to data-intensive business is the lack of coherence between different data sets. Ren et al. (2020) describe this as the gap between human- and machine-actionable search, access, integration and analysis. The FAIR principles should constringe this gap through the improvement of the findability and accessibility of data sets. They highlight the possibility of prioritizing interoperability through the FAIR principles. Currently, systems often have different data types, platforms, and formats and the systems being used are of a heterogeneous nature and variable. Data might be described in an ambiguous way due to a lack of accurate metadata. Lastly, it is mentioned that data is often described in a way that is either machine readable, such as JSON or XML, or human readable, such as HTML.

Wolf et al. (2021) take a more technical perspective. They call for a new approach to scientific workflows, specifically the construction, deployment and evolution of the software related to the workflows and the data that is used within. Workflows are consistent and present in every line of work, yet they are difficult to define. They state that the implementation of the FAIR principles can greatly benefit the reusability of the workflows every organisation already works with. However, they do note that the definition of reusable might depend on whom you ask. Next to this, they argue that implementation of the FAIR principles could aid organisations in defeating their technical debt. Technical debt is defined as the degree of effort a person must put into reusing a set of data (or code). Every piece of information that is not contained in a (meta)data set incurs technical debt. As FAIR aims to create detailed (meta) data sets with as much information as possible, it could help organisations to lower their technical debt.

2.2.2 Requirements

Unique & persistent identifiers

One of the goals of the FAIR principles is to make data easier to find. Findability can be increased through the use of descriptive metadata, which can aid in discovering and verifying data sets more easily. This is supported by Hauschke, Nazarovets, Altemeier, & Kaliuzhna (2021), who go as far as stating that metadata is the key to

enabling findability. The descriptive metadata can be paired with a unique identifier to not only help to make a data set more findable but also more accessible through the use of a data ecosystem. It will be important for an enterprise to decide on one structured way to identify data objects. An example of how this is done in research is the digital object identifier or DOI. It is important to note that the metadata, too, requires identifiers that are unique and resolvable, as well as detailed provenance information so that discoverability is increased (Juty et al., 2020; Garcia et al., 2020). Hauschke, Nazarovets, Altemeier, & Kaliuzhna (2021) highlight the notion that different aspects of information require different identifiers, as this provides different entry points. Additionally, identifiers must not only be unique but also persistent. This is because persistent identifiers can aid in correctly identifying an object. Unique identifiers ensure that every bit of data has its own identifier. This means that identical pieces of data in different data sets will have different identifiers. The unique identifier is only used to refer to the current location of the data object, meaning it can change. Persistent identifiers are a long-lasting reference to a data object. The persistent identifier will always be the same, independent of the location of the data object. In order to make a data object easy to find, it will need an identifier that is both unique and persistent (Muilenburg, 2021).

FAIRification

In order to implement FAIR guiding principles in the process of data storage and retrieval, it is important that the data itself is formatted in such a way that the FAIR principles can actually be implemented. Jacobsen, Kaliyaperumal, et al. (2020) refers to this process as "data FAIRification". They illustrate a workflow that consists of three parts: pre-FAIRification, FAIRification itself, and post-FAIRification. Below, a figure can be found with the step by step description as created by Jacobsen, Kaliyaperumal, et al. (2020) (figure 2.3). While they might not provide a full framework for implementation, they do provide a plan that incorporates a crucial part of the implementation process, being the preparation of data. As mentioned previously, the process is split up into three parts. The three parts will be described in more detail below, as the specifications are important to ensure the process is performed correctly.

Pre-FAIRification
The pre-FAIRification phase consists of three parts. First, the objective for the FAIRification must be determined. This could be a goal such as "increasing interoperability through improved metadata". Important here is that the data is available and accessible. Secondly, it is important that participants have knowledge of the data set and the FAIR principles. The objectives can be set by a variety of external stakeholders, as well as requirements from internal sources. It is recommended that participants start with a small subset of the data, and later iterate several times to incorporate the entire data set. Next, the data must be analysed to see what needs to be done. This could include steps such as identifying the formats of the data or investigating if the data already is FAIR in some way (an example of this would be the inclusion of metadata or certain unique identifiers). The final step of the pre-FAIRification process is to analyse the metadata. Important to note is that in this phase, the metadata does not have to be interoperable yet, as this is part of the FAIRification process in the next phase. The analysis of the metadata could include

actions such as seeing if there are metadata and whether it describes the correct data or aligning whether the metadata is rich enough through detailed provenance. Musen (2019) remarks that this process is often challenging. A tool that fills out metadata templates could be useful here. This is because users often experience the creation of a comprehensive and well-defined meta-model to be demanding. A tool that would make the task less bothersome, would be a good addition to motivate project teams to go FAIR. Additionally, they believe that a set template for metadata can help enforce comprehensiveness in the contents of the metadata. They expect that these templates will lead to more reusable and interoperable metadata, as the details should be similar due to the use of the same template (Jacobsen, Kaliyaperumal, et al., 2020).

FAIRification
According to Jacobsen, Kaliyaperumal, et al. (2020), this phase consists of 3 steps, where two out of three are performed for both the data and the metadata. First, a semantic model must be defined for both the data and the metadata. These models can be viewed as a guide for the next steps in the process. As the FAIR principles occur in more and more data sets, it is important to review whether a semantic model exists already and if reuse is possible. In case there is no semantic model available, a new one must be created. This can be done in three steps: build a conceptual model, define ontology terms, and combine these two to create the semantic model. Ren et al. (2020) underline the importance of a well-defined ontology and mention that a well-defined ontology provides a new way of expressing knowledge of a concept, where new knowledge can be inferred easily. The use of a well-defined ontology can greatly improve the interoperability of a resource. S. R. Wilkinson et al. (2022) add to this by saying that the difficulty of the data can be a challenge in mapping the data to a model. They suggest that participants translate and label the data in such a way that it is understandable for humans and machines, while also using FAIR vocabularies.

Participants must keep in mind that the semantic model must be readable for machines as this could greatly influence the effectiveness of the semantic model. Additionally, it is important that the semantic model makes a clear distinction between the data itself and its types. Next, the data and metadata must be made linkable. This is described to be a challenging part of the process, as the to-be-used method is case-dependent. It is of significant importance that the representation framework is understandable for machines and that the semantic model is accurately associated with the data and metadata. This is of importance because it enables the data sets to be available for future applications and enhances possibilities for scalable interoperability. The final step of this phase is to make the data ready for consumption. This means that it has to be made available to machines and humans through a variety of sources. A potentially great option would be to deploy it using a FAIR Data Point (FDP), which allows for controlled access and assists in ensuring that the data set is human- and machine-readable. Another author underlines the importance of the integration metadata and goes as far as mentioning that a Resource Description Framework format is required to ensure that the metadata is consistent with the ontology (Hauschke et al., 2021).

Post-FAIRification
The final step in the iterative process of FAIRification is to assess whether the created data set is FAIR. This can be done through an evaluation of whether the objectives that were set in the first phase have been accomplished, or whether the data set passes the FAIR status using a FAIRness assessment tool.

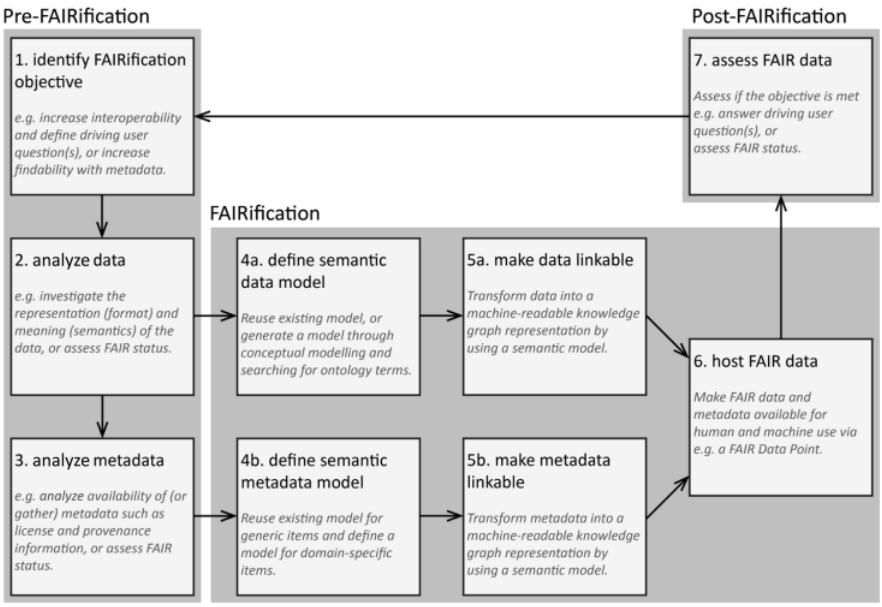

Figure 2.3: FAIRification process as described by Jacobsen, Kaliyaperumal, et. al. (2020) p. 58

The process of FAIRification should ensure that the data is of a high enough quality, which is a requirement mentioned by Hauschke, Nazarovets, Altemeier, & Kaliuzhna (2021).

Garcia et al. (2020) articulate ten rules to make data FAIR, although the process is not referred to as "FAIRification". Some of the stated rules are equivalent to the process that Jacobsen, Kaliyaperumal, et al. (2020) describe, but some rules are unique. The most important unique rule that is identified, is the definition of access rules. It is suggested that organisations make the decision between making it open to every actor who has an account or making access limited through an access-request mechanism.

Other requirements

In order to make the data accessible for humans and machines, it must be readable for both of those groups. This can be done through the use of an open format with standards and protocols. Hauschke, Nazarovets, Altemeier, & Kaliuzhna (2021) suggest using CERIF. Next to this, the use of W3C standards is suggested to enhance syntactic interoperability. Semantic interoperability can be ensured with the use of the same format, CERIF. This requirement is supported by Garcia et al. (2020), who suggest using various types of formats for varying types of data.

Differences in requirements for public and private organisations

An overview of the differences in the FAIR principles between a research context and an enterprise context is also given. Interestingly, in both contexts, the Reusable principle is the same in terms of the need for rich documentation. The other principles, however, have significant differences between them. Findable in a research context refers to certain search functions within repositories, whereas in an enterprise context, it refers more to the presence of the data assets in an inventory. Secondly, accessible in an enterprise context refers more to an interface and platform that shows different data layers, while taking into account the need for strict access rules and approval processes for sensitive data. This should encourage employees to use the data in more sophisticated ways. This is similar to the research context, where the focus is mostly on the identification of users who have access to sensitive data. Lastly, interoperable has rather different meanings in the two contexts. In research, it highlights the use of standardized formats to ensure that datasets are usable across different platforms. In an enterprise context, the focus lies more on the quality of the data and its references to other data, as enterprises often work in the same ecosystem, thereby decreasing the need to be interoperable with other systems outside of the enterprise (Labadie et al., 2020).

2.2.3 Management of implementation & data

In the previous section, the data FAIRification process was explained extensively. Additionally, Thompson, Burger, Kaliyaperumal, Roos, & Santos (2020) suggested the use of a DMP tool, as this is a process that might require a lot of time and effort, it is important that the FAIRification process is included in any Data Management Plans (DMP) that an organisation might have. This is so that the budget can be assessed properly so that it will not obstruct implementation (Jacobsen, Kaliyaperumal, et al., 2020).

S. R. Wilkinson et al. (2022) describe the FAIR principles to be of a non-prescriptive nature. This means that it was a deliberate choice not to add a checklist for implementation and striving for compliance. They also highlight the notion that some of the guidelines and principles will be significantly easier to incorporate.

Pană, Ivanoaica, Raportaru, Băran, & Nicolin (2021) have implemented FAIR into a database for geological information. They describe their process as challenging, yet rewarding. They have encountered that data standardization is a crucial step in the process of going FAIR, as this eases the implementation of the principles to data sets. To this end, they used data ingestion pipelines that collected, parsed, and extracted necessary data sets. To make the process more straightforward, a database format that was easily interoperable and reusable due to clear metadata identification was chosen. Then, the researchers developed authentication methods that made the data findable and accessible based on username. They refer to their new system as a "local FAIR environment" (p. 4), but indicate that the local environment will be challenging to incorporate with other organisations outside of their research group.

Roszkiewicz (2010) calls for an increased focus on metadata management. It is claimed that the management of metadata increases productivity, improves decision-making, and eases compliance with certain regulations. Additionally, it should increase the options for scalability, which could be important as the amount of data seems to be ever-increasing. Currently, enterprises do not take their metadata seriously. Departments that created metadata models did so separately from other departments. The best strategy for an enterprise there would be to combine these models into one enterprise-wide metamodel. This model can then be integrated with the ERP system and digital asset repositories.

A possible way for enterprises to incorporate the FAIR principles in their data is through the use of a data catalogue, which is an example of a digital asset repository. The role of a data catalogue can be defined as "maintaining a register of digital assets (such as data) through the detection, description, and management of databases" (Labadie et al., 2020). Russom (2017) adds to this that data catalogues aid enterprises increase their focus on data and become more data-driven. The data catalogue is explained in more detail in Appendix C. This section has presented several requirements for the implementation of the FAIR principles and provides suggestions for adherence to these requirements. The next section provides tools that can aid enterprises in their use and implementation of the FAIR guiding principles.

Tools for management of FAIR

Thompson, Burger, Kaliyaperumal, Roos, & Santos (2020) suggest a variety of tools to support the management of data and the FAIRification process, as they recognize that an ecosystem of tools and standards for the support of FAIRification and the consumption of FAIR data is lacking. Additionally, they hear the need for an answer to the question of how organisations can become more FAIR. A large number of principles describe the goal very well, however, the route to that goal is left unspecified. Next to this, certain principles require considerable infrastructural changes like communication protocols or search engines. Lastly, some principles demand a consensus or standard, even though the principles themselves are not a standard. Certain terms like "open" and "rich" call for a clear definition, while these are not yet provided.

The first category of necessary tools has a relation to the management of FAIR data and the planning thereof. A Data Management Plan (DMP) tool should assist researchers in the creation and maintenance of their DMPs. Some examples of tools that serve this purpose are the Data Stewardship Wizard (DSW) and DMPOnline. The latter has been rapidly adopted by researchers and organisations alike as their standard tool for the production of DMPS. The first tool, the DSW, consists of a knowledge model that guides users in the production of the Data Management Plan by taking them by the hand during the process. Importantly, it also allows the user to create FAIRness metrics to assess the FAIRness of the data they produce. It is suggested that there will be future attempts at machine-actionable DMP tooling, thereby increasing interoperability. Additionally, the remark is made that DMP and FAIR metrics tools might co-evolve in the future (Jacobsen, de Miranda Azevedo, et al., 2020).

Once there is a plan for the management of FAIR data, the next step is to create the FAIR data for the system. This can be challenging as the FAIR principles require a lot of attributes with association to the data, such as metadata, identifiers, and provenance. There is a variety of options for a "FAIRifier" tool, but all of them have different use cases. Some perform the process at the source, whereas others do so after the data has been obtained. A selection of technologies, such as Linked Data, can be incorporated into the FAIRification process, and it will depend on the use case which ones are most beneficial. Finally, some tools support automated workflows, whereas others can only perform the process through manual input. A general FAIRifier has been produced which has its basis in the OpenRefine tool for data cleaning and wrangling. Nevertheless, it must be noted that this tool will not be usable for every organisation or use case(Jacobsen, Kaliyaperumal, et al., 2020).

After producing FAIR data, the data must be published on a platform where users can extract the data. This is where the challenge of accessibility comes in.: who is allowed to see what? Additionally, this is likely where the technical challenges arise. The datasets must be machine-actionable all while adhering to organisational, legal, and regulatory procedures. The infrastructure will play an important role in this, as data consumers should be able to get the data in a predictable way. The use of a Fair Data Point (FDP) is suggested as its benefits are twofold. On the one hand, it will provide compliance with the FAIR principles, while on the other hand providing an easy-to-use platform that likely already has access to repositories (da Silva Santos et al., 2023).

Finally, the last tool that could be needed is related to the final principle: reuse. The misconception exists that data assets will always be reusable if they adhere to the other three principles. However, this is not necessarily the case. The fulfilment of the last principle greatly depends on the data being rich in accurate attributes. This principle does not necessarily require a tool, but more so a standard for metadata and the registration thereof. Currently, the FAIRsharing registry exists, but it does not contain every data set that exists and is unlikely to do so in the future (Jacobsen, de Miranda Azevedo, et al., 2020).

Jacobsen, de Miranda Azevedo, et al. (2020) composed an extensive list of implementation considerations, which can be seen visualised in figure 2.4. As seen in the figure, many of the principles have certain considerations that need to be considered when implementing them.

FAIR	*Principle*	*Implementation*
Findable	F1	Organisations must think about the longevity of the identifiers they plan on using. These identifiers should outlive the termination of a project or department. A persistent identifier might require a third party to provide them.
	F2	When defining the necessities of metadata, the richness should be properly described so that it fulfils its purpose.
	F3	The created metadata model must be machine-actionable and has to be linked to a resource
	F4	Organisations must think about how they design their search structure. The goal is to create a search strategy that is as uniform as possible so that no results are overlooked.
Accessible	A1	-
	A1.1	Access protocols must be documented as explicitly as possible. The protocols must be part of the machine-readable metadata.
	A1.2	Organisations must decide whether metadata should be freely available, even if the original data is not. Additionally, the decision must be as generic as possible to prevent confusion.
	A2	Organisations must consider what policy they use for metadata once the data is unavailable. This decision must be present in the documentation and a machine-actionable scheme.
Interoperable	I1	Organisations must define a procedure for communication in which the language is explicitly mentioned. The data must be interpretable in the chosen language, and it is key that the language is understandable for generic machine agents. This can be achieved by making the data available in various formats.
	I2	It is important that the structures used themselves are also FAIR. This means that ontologies and semantics have to be available, and thesauri that are not universally accessible should be changed.
	I3	Relationships between data and metadata, or metadata among itself must be explicitly referenced and an appropriate relationship should be chosen according to the FAIR principles.
Reusable	R1	-
	R1.1	Organisations must find a way to divide between a license that allows access to the data and a license that allows access to the metadata. Generally, organisations should document what the license requirements are.
	R1.2	Organisations must attempt to provide as much provenance as possible. Additionally, they must define a method that decides what provenance is useful and what is not.
	R1.3	Organisations must define their standards and how data and metadata can adhere to them. Additionally, it must be represented using a more global standard, so that the (meta)data remains findable for third-party agents.

Figure 2.4: Implementation considerations for the FAIR principles as given by Jacobsen

The expectation is that more tools will emerge as the FAIR principles gain a reputation. This means that the process of FAIRification and everything else that comes with being FAIR might become easier in the future.

The implementation of a new system or methodology often asks a lot from em-

ployees. Not only will things change on a technical or regulatory level, but a change in mindset and company culture is often also required. This change must be handled with care as it is a large factor contributing to an implementation's success. The next section will focus on change management and what empirical research has shown about how it can be managed successfully.

2.2.4 Change management

The implementation of a new method or system often causes uproars from employees (Bharathi & Mandal, 2015; Tome, 2014). Research has been performed on the best practices of change management and on frameworks that shape the route to change management. The framework that will be developed will serve as the preparation for adoption. The preparation of employees is part of the adoption of a new system or way of working, as it is said to aid in the acceptance of the adoption (Akrong et al., 2022; Shah et al., 2011; Amoako-Gyampah & Salam, 2004). This is why the decision was made to add a change management plan to the framework.

Best practices

It is important that the level of engagement and involvement from (at least a part of) employees is high. This is where employees can voice their concerns and wishes and enables the organisation to communicate effectively with all of their departments (Akrong et al., 2022; Bharathi et al., 2012; Amoako-Gyampah & Salam, 2004; Mesicek et al., 2021; Shah et al., 2011). Next to this, proper training for employees who require this is crucial to the success of an implementation, but also employee success afterwards. People need to understand how a system or methodology works and why it is necessary before they can effectively work in it in the most efficient way possible, therefore this is a critical factor in making a new method work (Supramaniam & Kuppusamy, 2011; Tarhini et al., 2015; Alballaa & Al-Mudimigh, 2011; Van Hau & Kuzic, 2010).

The creation of a well-functioning project team is important. The use of a so-called "project champion" can aid in this. Such a person ensures that everyone in the team stays on track and is still on board with what is going on (Supramaniam & Kuppusamy, 2011; Alballaa & Al-Mudimigh, 2011; Shah et al., 2011; Van Hau & Kuzic, 2010). Top management plays an important role in such projects, and having them on board has significant importance in succeeding (Masheshwar & Javalagi, 2019; Tarhini et al., 2015; Akrong et al., 2022). They can ensure company wide support, which is another important aspect of large projects like this. If the largest part of the company has an understanding of why such a project is important and why it must happen now, they are likely to be more lenient when something does not go according to plan (Tarhini et al., 2015; Shah et al., 2011). [1]

[1] Please note that some of the content of this section was adapted from an earlier report by the same author.

Frameworks & models for change management

Two frameworks and one model for change management have been selected for further research as these are often used in an enterprise context. The first one is the Mckinsey 7S and the second one is Kotter's model for change. A brief explanation will be provided in this chapter but the two models will be described in more detail in Appendix C.

McKinsey's 7S model is a conceptual framework that aids organisations in diagnosing the causes of their organisational problems and leading them towards improvement. The framework splits seven elements into two categories: hard elements and soft elements. The seven elements are strategy, system, structure, shared values, skill, staff, and style (Waterman et al., 1980; Waterman, 1982; Channon & Caldart, 2015). In Appendix C, more information about the framework and its elements can be found.

Kotter's models for leading change & accelerating change is the second framework related to change management. In figure 2.5, the differences can be found. Next to the steps, guidelines and actions are specified as well. The first step starts by creating an environment (or ecosystem) for change. Then, steps 2 until 6 all serve the purpose of preparing the organisation for the change. Lastly, steps 7 and 8 serve the purpose of leveraging the successes thus far and institutionalising the change to make it permanent (J. P. Kotter, 1995, 1996; Kotter, 2012). For more information on the steps and related actions, see Appendix C.

Step	Leading change 1996	Accelerating change 2014
1.	Establishing a sense of urgency	Create a sense of urgency
2.	Creating a guiding coalition	Build a guiding coalition
3.	Developing a vision and strategy	Form a strategic vision and initiatives
4.	Communicating the change vision	Enlist a volunteer army
5.	Empowering a broad base action	Enable action by removing barriers
6.	Generating short-term wins	Generate short-term wins
7.	Consolidating gains and producing more change	Sustain acceleration
8.	Anchoring new approaches in the culture	Institute change

Figure 2.5: The different steps in both of Kotter's models

The final model for change management that was selected based on the advice of a validation participant is the ADKAR model. This model splits change management up into two zones: the enablement zone and the engagement zone. The methodology consists of five necessary outcomes that need to be achieved before a change will be successful: Awareness, Desire, Knowledge, Ability, and Reinforcement (Hiatt, 2006). More information about this framework can be read in Appendix C.

2.2.5 Reflection on theoretical background

To the best of the author's knowledge and the extent to which the review was performed, the literature review has shown that a lack of empirical evidence and validated research on the topic of FAIR in an enterprise context exists. This was

identified as the search terms "FAIR principles" AND "IT consultancy" yielded no relevant results in the utilised databases. It was noted that many enterprises do use the principles under a different name and without actively knowing that they are, but will not openly publish this information as it is a competitive advantage to them. This highlights the relevance of the thesis. Most papers have not conducted extensive analyses of the use of FAIR within the chosen sector. The literature, however, did provide clear insights on the use of FAIR in an academic context as well as on the way some enterprises organise their data through the use of a data catalogue. The papers showed elaborate strategies for the adoption of FAIR and little guidelines on implementation were presented, even within the research industry. Because of this, the literature has contributed foremost to assessing the state of research in this particular domain. Additionally, the author noted that the use of a data catalogue occurs more often, even with FAIR principles incorporated into it, but this goes unseen as they often have different names or structures. Lastly, it has become clear that certain management methods and tools for FAIR have been developed, but that there still does not exist an overview or procedure with steps for research organisations and enterprises to orchestrate their FAIR process. Although it is mentioned often that the FAIR data principles are not a standard (which is likely why there is no set list of steps for implementation), a methodology could help enterprises to understand the changes they need to make and the benefits they will gain from implementing the principles.

The empirical part of this research will attempt to answer the sub-questions with relation to the implementation of FAIR in IT consultancy to serve as the basis for the framework for the preparation of the adoption of FAIR in the consultancy sector. The literature review on change management did bring forward interesting success factors, as well as important pitfalls to look out for. These will be incorporated into the change management plan that will be a part of the final methodology.

Chapter 3

Results

This chapter contains the highlights from the case study. This includes quotes from the case study to support the findings. First, the participants will be introduced after which several parts of the case study will be discussed in more detail. The parts are the current system, FAIR, and Change management. Additionally, the draft framework will be presented. An explanation of the title of the framework will be provided, accompanied by the highlighting principles and the change management guide.

3.1 Case study results

The case study was performed at one enterprise with managers and senior managers from several departments that all have a connection to IT. First, the participants will be summarised. After that, the results that are most important to the development of the framework will be discussed.

3.1.1 Summary of participants

In this section, a short, anonymous description of the experts will be given. All respondents work for the same organisation in The Netherlands. Their identities have been made anonymous to prevent the possibility of tracing back to the company they work for. The details can be found in table 3.1. The participants have a varying amount of experience within the enterprise the case study was performed at.

Table 3.1: Case study participants

Participant	Job title	Department	Experience at the enterprise
Participant 1	Senior Manager	IT audit	9 years
Participant 2	Senior Manager	Advisory	16 years
Participant 3	Manager	Advisory	7 years
Participant 4	Senior Manager	Advisory	10 months

The full overview of all answers related to benefits, challenges, needs, and requirements can be found in table D.1. The answers to questions about change management can be found in table E.1. Relevant quotes from the case study were extracted from the transcripts and can be found in Appendix F.

3.1.2 Current system

Several case study participants mentioned that many employees store data locally on their laptops or desktop computers and that this leads to challenges in completing data sets. As the local machines are not accessible to others, this creates disparities in the data that is available to team members. One participant stated that there is a policy related to this topic, but that many employees are not inclined to adhere to the policy. Additionally, they voiced a desire for more employees to be involved in the process of data storage so it becomes a shared responsibility. Currently, there is a preference for team-wide use of Microsoft tools, such as Microsoft Teams as this is accessible to team members. On these platforms, several people can work on one document at the same time, which is convenient for version control, which is something one participant found desirable.

When the participants were asked about their opinion about the current system and its policy, the answers were diverse. One participant expressed that they found the guidelines clear and easy to follow, whereas the policy was not clear enough for another participant, who preferred a checklist over "vague" guidelines. It was highlighted that the enterprise provided training on the policy and method for storage and archiving, but participants doubted whether those responsible were up-to-date with everything. One participant explained that the guidelines related to iManage were followed, but that their team developed their own method on top of that to organise their files. IManage is the system that this particular enterprise uses for the storage of its engagement data.

Additionally, it was noted that the policy for data storage in iManage starts too late compared to the start of the engagement. Because of this, those responsible get behind on their task of storing files in a certain way, which leads to the tendency of doing everything at the end, thereby risking the loss of data during the engagement. A requirement was voiced about the new system, which states that the information necessary for data storage must be provided earlier in an engagement than what is currently the case.

Lastly, participants were requested to express benefits from the current system. All participants agreed that although the system is not perfect, it is great that the entire enterprise uses the same system. According to one participant, the current method prevents data breaches because they try to delete as much sensitive data as possible, as they mention:

> *"In the case that somebody enters the system, at least there are no privacy-sensitive files in there."*

Most participants mentioned that there is a large focus on confidentiality and security and that it is something that should never be overlooked in the design of a new system.

Finding stored data in the current system is not an efficient process. One participant mentioned that they thought that keywords or tags would be extremely beneficial in this case, and that they would like the system to have a search function such as Google's. Another participant points out that they would like to see an

increase in focus on metadata, as this, according to them, will make the process of finding and reusing data more practical.

3.1.3 FAIR

After investigating the current system and requirements, the case study proposed the solution of the introduction of the FAIR guiding principles. Participants agreed that the principles are clear, with the exception of "interoperable". Most participants agreed that the principles are definitely useful in the field of research and think that sharing data with other institutions in that industry will lead to momentum in research. One participant was shocked that these principles were necessary, as technology has advanced astonishingly over the past few years. They remark:

> *"We move to Artificial intelligence. We have this chatGPT. It's crazy what is out there. And then you think we still have to think about finding stuff. Is it accessible? Is it reusable? It's really strange, but apparently, it is needed."*

Thereby voicing slight disappointment in the enterprise for not being as up-to-date in this regards as they would like to see. Participants mention that within the enterprise and her daughter-firms data is shared often but that this is not an orchestrated process. This is where the benefit of FAIR is first acknowledged. When asked about sharing data with their competitors, all participants said that they believe it is of high importance, but that the current competitive atmosphere makes this impossible. One participant strikingly remarked that the ecosystem of the industry must first change from competitive to cooperative before sharing across firms would be a possibility, thereby not ruling the option out for the distant future. The FAIR principles could aid organisations to structure their data storage approach, according to several participants. Currently, large amounts of data are stored in silos and there is no centralized way to store data sets.

When asked about the potential of the FAIR principles, one participant remarked:

> *I think it's very clear and very, It's actually pretty logical and summarized in those principles. If I think about data, maybe interoperable would be the last one I would think about, to be honest, but I think accessibility findability and reusability are something that also comes to mind.*

Another participant agreed with the aforementioned statement, but added the following:

> *I can't disagree with this. Of course, some of them are a little bit open door and that is the case in many models. They describe sort of the ideal situation, but it takes a lot of maintenance and ownership and discussion. To get to that point.*

They expressed interest in the FAIR principles but mentioned that they are concerned that it will be too much maintenance and that employees will not take responsibility for this.

When discussing the potential implementation, all participants expressed the need for a change management plan to guide the organisational change required before the principles can be used to their optimum form. Next to the required industry-wide mindset shift, several participants stated that a shift in enterprise-wide mindset was also required. Currently, some employees still consider data to be a powerful asset and might not be willing to share this with others.

When asked what principles they thought were most important, the same principles were mentioned often. Participants were most appealed to foundational principles findable, accessible, and reusable. The principles that were mentioned are: F1, F3, F4, A1, A1.2, R1, and R1.3. This is likely because most participants mentioned that they thought that these principles were the clearest and easy to visualise. Principles related to the openness of the data source led to confusion for some participants, highlighting the requirement that the new framework must clearly explain that open does not mean accessible to all.

When considering additional principles to be added, almost all participants highlighted the notion once more that they would prefer the system to be internal at first. Additionally, all participants also stated that it worried them that there was no principle related to the quality of the data and the way metadata was utilized for version control and maintenance. A possible addition, therefore, could be a principle that focuses on maintenance, updating, and quality assurance of the data. Participants believed that this would make the data even stronger. In order to do their tasks well, they must be able to trust that the data they work with is the latest version and that it is correct.

The participants were also asked in what way they would like to see the (added) principles presented. One participant mentioned that they would prefer a checklist, whereas another participant remarked that a framework would be their preferred way.

> *A framework which is also easy to understand, which with some principles you can always say keep in in the back of your pocket or in the back of your head to to start thinking actively. How to operate and manage your data.*

3.1.4 Change management

The final part of the case study was related to change management. Change management was already identified as a crucial step in the implementation of a new process or system, which is something that case study participants all agreed with. In the past, not all changes in the enterprise were managed as well as participants would have liked. They mention that the addition of training and workshops are important to educate the necessary skills to employees. Additionally, a fondness for the use of workshops, pilots, and focus groups was expressed as this incorporates employees into the implementation process and helps them become a part of the solution. One participant added to this that these activities can help create small achievements that can be leveraged later on. Several participants remarked that pilots and workshops can help create change ambassadors. These are people that can be your air to other employees. This can help speed up the process of spreading

the information across the enterprise, using word-of-mouth advertising. Secondly, change ambassadors can help advocate for the solution and help the implementation team leverage the achievements thus far. One participant added that the change ambassadors should not disappear once the implementation is finished. That is the moment where they are needed the most because then people actually have to work in the new system. It is beneficial to have a person available for employees to go to when there are questions or concerns.

During the discussion on change management, the topic of employee (or enterprise) mindset made a reappearance. One participant stated that an ecosystem for change and a mindset for change must be developed within departments before the change even starts. They mentioned that it is difficult to make culture and mindset a tangible parameter while this is important in the acceptance of the change. This is why they think that a certain change in mindset will have to be developed before anything else changes, although they made the additional remark that this can also be a gradual process once the change is implemented.

One participant voiced a strong need for long-term focus as this is something they believe is currently missing. The way of working often changes and the systems change with it. This makes it challenging to convince people to adhere to guidelines, as those guidelines will likely change within a short time frame (e.g. a few years). A long-term focus would greatly help the acceptance of the new framework.

3.1.5 Conclusion

The case study has shown that (senior)managers have a diverse vision about the current challenges. However, they seemed to agree that the FAIR principles can aid in the challenges they experience. When asked about FAIR, most of them agree that the principles are clear, although some experienced difficulty with the specific definition of some of the principles. All participants see a benefit in implementing the FAIR principles to some degree, but there is no coherent answer as to what that degree should be. All participants mentioned that it is still too early for data sharing with competitors, so the FAIR principles will first have to be implemented internally. Perhaps at a later time, they can be rolled out to other enterprises and one large data ecosystem can flourish. When asked what they were missing in the FAIR principles, all of the answers were related to the trustworthiness and quality of the data. This is a factor that could be added to the framework that incorporates the worries voiced by the experts. The advice for change management was mostly based on experiences they had themselves within the enterprise, which is relevant information because it shows where pitfalls and gaps have been in the past. The next section contains the FAIR-iT framework, which is based on a combination between the literature and the results from the case study.

To relate the results to Wieringa (2014), a small overview of the goals of the treatment design will be discussed. The case study has shown specific requirements that will aid in achieving the goal of improving the data storage process within enterprises, as stated by participants. The FAIR principles are a treatment that is available but does not cover all requirements for the enterprise the case study was performed at. This is why the new treatment (or artefact) of the FAIR-iT

framework was created. The new treatment should cover more of the requirements and therefore contribute better to the set goal. The specific requirements that were extracted from the case study can be found in Appendix E.

3.2 Introducing the FAIR-iT framework for enterprise preparation for the adoption of the FAIR principles

To place the FAIR principles in an enterprise context, certain adaptions will have to be made as concluded from the case study. The proposed framework will be based on literature and case study results, even though the literature provided little information about what the FAIR principles would do in an enterprise environment. Since the literature review shows little research on what the FAIR principles mean in an enterprise context, the opportunity for a synthesis between literature and case study results is insignificant. The literature will therefore serve as an inspiration for the framework that is developed and will be referenced where possible. This is not the case for literature on change management, so in section 3.2.4, a synthesis for that part of the literature review can be found. The upcoming sections will show the FAIR-iT framework, explain it, and provide a description of how the principles are adapted to better fit an enterprise context. Lastly, an accompanying change management guide is given which focuses on steps to undertake to attempt a change journey with as little resistance as possible. The created framework applies to the process of data storage of engagement files, as several participants highlighted that this is something that their company and teams can improve on.

3.2.1 Why FAIR-iT?

Considering the needs that were voiced by participants in the cast study, the decision was made to add two extra letters to emphasise enterprise focus points. The "i" originates in the notion that all participants mentioned that their sector is not ready for intensive data sharing with other companies and that it is preferred if the data that an enterprise has, stay within that enterprise. However, as one participant noted: Perhaps in the future, there is more room for collaboration and cooperation, rather than competition. This is why the choice is made to make the second "i" a lowercase letter. This is something that might alter in the future, it is not a constant. For now, the "i" stands for "*internal*". The principles that will be shown later, will refer to the lowercase letter as well to avoid confusion with the I from interoperable.

Then, the second letter added is the "T", which stands for *trustable*. Considering the results from the case study, all participants voiced concerns about data not being maintained, not being up-to-date, not being of high quality, or being manipulated. All of these concerns fall under the same category of *trust*. The users must be able to trust their data before they can optimize their use of said data. As one participant noted: the data users must become confident in and comfortable with their data. This is likely to happen if the data they are presented with is trustworthy and accurate. Trust is also created through transparency, which is something that Inau,

Sack, Waltemath, & Zeleke (2021) explained as a result of the implementation of the FAIR principles.

3.2.2 The highlighting principles

The FAIR-iT framework has six principles. The four that were already defined in FAIR, will stay the same. The two added principles also have sub-principles and have as a goal to highlight necessities for enterprises. These are described below in table 3.2. Important to note is that the added principles do not necessarily provide large portions of new requirements. Rather, they highlight aspects of the existing four principles. The addition of the highlights on data quality and security makes the 4 original principles more applicable to an enterprise context as they address worries expressed by case study participants. This means that the FAIR-iT principles consist of four foundational principles and two additional principles for emphasis for applicability to the context at hand. The two additional principles highlight factors of importance to the governance of the data assets, rather than only archiving, finding, and sharing them.

Table 3.2: The highlighting iT principles

Principle	Number	Definition
i		
	i1	(Meta)data is shared within an enclosed data ecosystem, such as a data catalogue
	i2	(Meta)data cannot be removed from the enclosed data ecosystem by a third party without a request as contracted
T		
	T1	(Meta)data is actively maintained through frequent updates
	T2	(Meta)data quality is guarded requirements according to enterprise standards
	T3	Metadata are updated when the data is adjusted

The iT principles could be seen in the table above. Here, they will be explained in more detail. Starting with i1: (Meta)data is shared within an enclosed data ecosystem, such as a data catalogue. As already mentioned in chapter 2.2, enterprises often have a data catalogue in place rather than a database. The benefit of having a catalogue is that the data inside of it is already managed in an inventory. If tags and identifiers are added to this inventory, then a large challenge of implementing FAIR-iT has been taken already. i2 is about third parties removing data from the data ecosystem. This can be important to enterprises because it will occasionally happen that on a project, a third party is added to the contract to take care of certain parts of the work. This third party will then need access to files from both the client and the cooperator, the enterprise in this case. However, as the files are stored for internal use, a third party is not able to easily access and retrieve this data. How this is orchestrated, should be clearly explained in the contract that is made for the engagement. Additionally, a portal can be created for third-party data requests. Then the request can be reviewed and access can be granted if the request is approved and within the terms of the contract.

T1 states: (Meta)data is actively maintained through frequent updates. As one of the participants mentioned, a lot of the storage of the project data happens in the final few weeks of the project. This is not the desired situation, as it means that during the engagement, the files are not accessible to other employees. It will be important that (senior)managers frequently ask their project team members to send updated versions of the files they are working on. This will require a change

in mindset from departments as a whole as it is significantly different from the way the work is currently organised. As one case study participant noted:

> *"We haven't turned around our way of working to already use the system from the start [iManage]. More junior colleagues don't work in iManage. They work in the project system of our clients or the project system of what we use with teams and it's the engagement managers, and engagement partners' responsibility to make sure that all the documents are also properly stored in iManage. Whilst you would rather have that the entire project team works out of the iManaged application from the start."*

This quote indicates two important factors. 1) The current way of working is not the most efficient as not every member of the team has access to the iManage, which is the system where engagement files are stored. And 2) The previous implementation of a new system (iManage) perhaps was not managed in the most effective way as the way of working has not changed as opposed to the system that was active before iManage, while that is something that was required. This principle, therefore, will be the easiest executable if the access rights to iManage change from, for example, role-based to project-based. Then, once an engagement is considered finished, the access rules might change again back to role-based, so that not all documents are accessible to everyone in the company. The second principle, T2, refers to (meta)data being guarded to ensure high quality according to standards set by the enterprise. This is important as several participants stated that they were afraid that the data would get manipulated and then lose its trustworthiness. Therefore, it is important that the metadata is updated when the data itself is updated, and that the updates to the data are within the standards that are set in the enterprise. Lastly, T2.1: "Metadata are updated when the data is adjusted" relates to the remark made above. This sub-principle should aid in motivating, or perhaps even obliging, employees to update their metadata, as this is currently not a high priority. However, the identifier- and tag system depends on the metadata being of high quality. If employees will not provide rich metadata, then the search is not likely to become easier.

The next section explains the method for the creation of the FAIR-iT data. It will provide an ambiguous method for employees to prepare their data to make it FAIR-iT, and it will show the process the data goes through when it is received from a client or third party. This is an essential step in the framework as this is where the principles are used.

3.2.3 Data preparation

The goal of the FAIR-iT principles and framework is to prepare the adoption of the FAIR principles in an enterprise environment so that the data from past engagements can be found back more easily through the addition of proper metadata and identifiers. The framework consists. First, the two additional principles mentioned in the previous section serve as the basis of the framework. These are then utilized in the model in figure G.1. Additionally, a change management plan will be described to aid enterprises in their preparation of the adoption of FAIR-iT. This can be found in section 3.2.4. The implementation of the FAIR-iT principles themselves is *not* part of the framework, as the framework is meant to serve as a preparation

for adoption. More about this can be read in sections 6.2 and 6.3.3.

As mentioned previously, a shift in mindset and work approach is necessary for the adoption of the FAIR-iT principles. Employees will be required to prepare their data so that the final file can be considered FAIR-iT. This process is based on the process as proposed by Jacobsen, Kaliyaperumal, et al. (2020). They provided a 7-step process for the FAIRification of data, as could be read in chapter 2.2 in figure 2.3. Interestingly, the GOFAIR organisation itself proposes a different method for the FAIRification of data. This will be explained below. They, too, propose seven steps but these differ significantly from the ones proposed by Jacobsen, Kaliyaperumal, et al. (2020). The proposed process can be seen in figure 3.1.

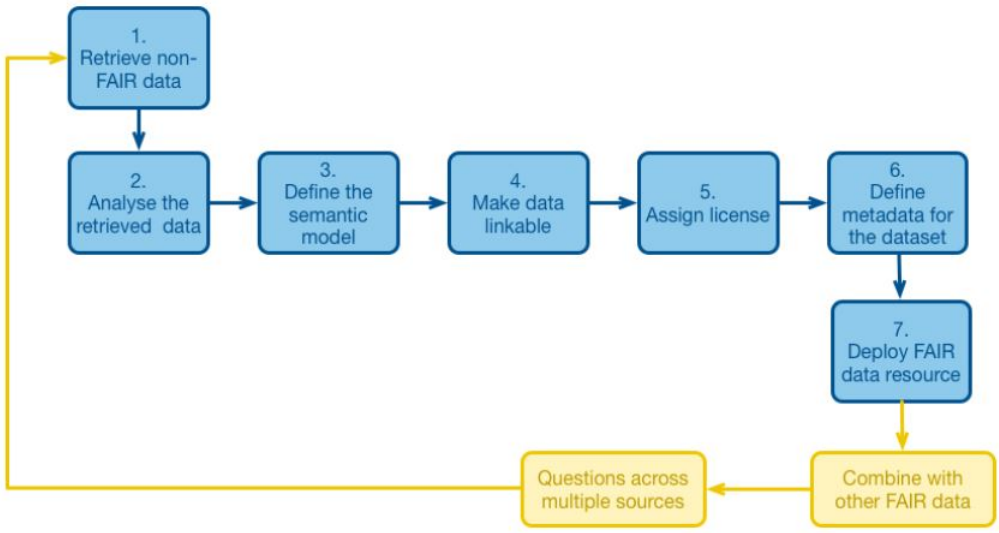

Figure 3.1: The FAIRification process proposed by GOFAIR

Important to note is that the model proposed by Jacobsen, Kaliyaperumal, et al. (2020) elaborates on the one proposed by GO-FAIR (2018). First, the differences between the two models will be identified. Then, a combination of the two models will be provided to guide employees to FAIR-iTify their data. The choice to combine the two models is based on the notion that both models have factors that are important for an enterprise context, whilst also both having steps that might not be as crucial.

The first identified difference occurs already in step 1. The model proposed by Jacobsen, Kaliyaperumal, et al. (2020) starts by identifying the objective. The model proposed by GO-FAIR (2018) starts by obtaining the data that is to be FAIRified. For an enterprise context, the latter would be more suitable as they often cannot choose what they receive. The next step is the same in both models. The goal of step 2 is to analyse the data. Then, Jacobsen, Kaliyaperumal, et al. (2020) define analysing the metadata as step three. GO-FAIR (2018) does not perform this step, and moves to defining the semantic model. Jacobsen, Kaliyaperumal, et al. (2020) also does this in step 4. The model created by GO-FAIR (2018) makes the data linkable in step 4, which Jacobsen, Kaliyaperumal, et al. (2020) does in step 5. Then, GO-FAIR (2018) assign the license in step five, which is something that

the process as proposed by Jacobsen, Kaliyaperumal, et al. (2020) skips completely. For GO-FAIR (2018) step six is where the metadata enters the data set, wheras the process suggested by Jacobsen, Kaliyaperumal, et al. (2020) already defined this in step 3. Their goals of step six is to host the FAIR data, which GO-FAIR (2018) does in step seven, although they refer to that process as "deploying the FAIR data". For Jacobsen, Kaliyaperumal, et al. (2020), step 7 is to assess whether the data is FAIR. In the table below (table 3.3), the processes can be seen for a clear overview.

Table 3.3: The processes as defined by Jacobsen, Kaliyaperumal, et. al. and GO-FAIR

Step	Jacobsen, Kaliyaperumal, et al. (2020)	GO-FAIR (2018)
Step 1	Identification of FAIR objective	Retrieve data
Step 2	Analysis of data	Analysis of data
Step 3	Analysis of metadata	Definition semantic model
Step 4	Definition of semantic models	Make linkable data
Step 5	Make linkable (meta)data	Assign license
Step 6	Make FAIR (meta)data available	Define metadata
Step 7	Assess if data is FAIR	Deploy FAIR data

Both processes have steps that are more important for enterprises than others. The process for FAIRification for enterprises will therefore consist of a combination of both processes with additions from the case study. For step 1, as mentioned previously, the step from GO-FAIR (2018) will be used. Step 2 is the same in both models. For step three, the step provided by Jacobsen, Kaliyaperumal, et al. (2020) will be used as it is important for the enterprise to research whether metadata already exists for that particular data set. Once this is done, the semantic model can be defined as proposed by Jacobsen, Kaliyaperumal, et al. (2020) (and GO-FAIR (2018) in step 3). With the semantic model defined, the data can be made linkable, which is step five. Then, a license should be assigned to define what users can access what data and who has the authorization to do what as step 6. GO-FAIR (2018) mention that the license should be part of the metadata. They also highlight the importance of a clear license as it increases the chances of the data being reused. The data set can be deployed or hosted on the data catalogue. This is step 7. The final step, making the process iterative, is to maintain and update the data once new data is provided by clients or third parties. The model for the creation of FAIR-iT data can be found in appendix G. As mentioned previously, the steps take the definitions from GO-FAIR (2018) and Jacobsen, Kaliyaperumal, et al. (2020). The data users can follow this process when creating FAIR-iT data sources. As this process gets integrated into the working approach, more FAIR-iT data sources will become available, thereby growing the FAIR-iT ecosystem. Hopefully, when enterprises are ready for data sharing with other enterprises, the data ecosystem can gain even more richness as data is provided from many different sources.

In the process, the use of metadata is specifically mentioned several times as some of the experts mentioned that metadata is often overlooked in their enterprise. The goal of mentioning it in the process model is to increase awareness around the

topic. Additionally, the goal is to move towards an environment where metadata and its use becomes known to an extent where it becomes a normal step to undertake. It was mentioned in the case study that this is currently not the case and that many employees are not aware of what metadata is or how they can use it to their advantage.

Currently, the responsibility for the data storage lies with the engagement managers as just one part of their extensive task list. One case study participant indicated that there is a desire to share this responsibility within the departments. The idea of the framework is to create more ownership of the data across the departments. Therefore, the responsibility for the data will shift from the engagement managers to the PMO (Project Management Ondersteuning). This role already has a large involvement with the data process. A large portion of the tasks of the PMO are related to communication about the project to stakeholders and keeping an overview of the data and files. This is why the task of FAIR-iT-fication will fit well within this role description. The PMO role in generally fulfilled by a consultant or senior consultant. Involving these titles in the data process will create a broader sense of data ownership across departments than when only management functions are responsible. In smaller projects, there usually is no PMO role. In this case, the task of FAIR-iT-fication and data storage remains with the project or engagement manager.

To clarify the new process to employees, a flow diagram for the FAIR-iT-fication is provided. It is important to remember that the consultants and senior consultants gain an important task. They will have to make the data FAIR-iT, as it is impossible to leave this crucial task to (senior) managers. Currently, the (senior) consultants do not come into contact with the process of data storage on a daily basis, other than putting files in a Microsoft Teams folder or SharePoint. Once the FAIR-iT principles are implemented, their tasks will be expanded and they, too, will have the responsibility of ensuring that data is filed with accuracy and high quality. In order to make the steps clear to all employees, but especially the PMO, the process was visualised. Below, the process steps will be discussed in more depth with accompanying suggestions on how the step can be performed. Because FAIR-iT is a framework, no specific guidance on the steps will be provided. This is to make it applicable to a larger spectrum of projects as not every project has the same structure and requirements.

3.2.4 The change management guide

For the creation of the change management plan, an expert from the change management department was asked about what they think are crucial factors in implementing change within their own organisations. Additionally, the participants in case study were asked about their opinion on how change is usually managed within their company, and how they think that process could improve. Together with findings in the literature, their answers served as the basis for the first version of the change management plan. Below, a list of decisive factors can be found. In the literature review, two change management models were discussed: McKinsey's 7S model and Kotter's model for leading change. Additional information about these frameworks can be found in chapter C. The decision has been made to follow the latter for the change management plan. This is because the organisation at

which this thesis is written already has a change management framework in place that is based on Kotter's. Due to confidentiality reasons, the specific name of the framework cannot be shared but the details will be described under a different name.

Frameworks used for guide

The model described by the enterprise itself consists of 5 phases and 4 accompanying workflows. They are similar to the ones mentioned by Kotter in table 2.5. The first step is to align leaders and state the strategic aims and scale of the change. This is where the goal is to make sure that all leaders understand the change and why it must happen. Examples of actions to undertake are proposals and leadership information sessions. It is crucial that the need of the change is mentioned explicitly. Then, the vision is communicated to employees, who will start to create ownership over the change. This is where you create department- or company-wide awareness. This can be done through an ambassador network, a clear communication plan, and a change story. Depending on the organisation, these will have very different contents and characteristics. Thirdly, the change vision is translated into clear indications of what the change means to employees and who it will affect. This is where employees are prepared to work with the changes. Here, an impact assessment is a useful tool to deploy. It is also important to inform employees about their (potential) new job roles and tasks. Then, the organisation is moved towards the state where employees can work with the new system or change. This is where employees are trained to work in the new system. This can be done through e-learning, gamification, or virtual classrooms. Additionally, this phase requires a lot of support and there should be change managers available to answer questions. The final phase focuses on the ability of organisations to sustain the change to make it permanent. A tool for this is a KPI dashboard. The four workflows focus on leadership on change, engagement and communication, the impact of the change, and the transition the employees must go through when adapting to change.

Synthesis of literature and expert opinion

In table H.1, the synthesis between the change management advice in literature and the change management needs from the case study are shown, which can be found in appendix H. All of the terms from the literature are explained in chapter 2.2. For some of the concepts obtained from the case study, an explanation was given in section F.3.

As mentioned previously, the implementation of the FAIR principles already will require a large shift in the mindset of employees. As one case study participant said, sharing has to become the norm and people have to get used to giving up their advantage through data. The added principles, and especially the Trustworthy principle will require not only a shift in mindset but also a shift in the way employees work. Added attention has to go towards the way they process their data, but also requires them to occasionally go back to old files to see if there is new information or maintenance required. Adapting a new mindset takes time and employees should not feel forced to adopt the mindset over a weekend.

Draft change management steps

The following steps have been identified for the to-be validated change management guide. The guide consists of five steps with smaller sub steps that were taken from Kotter (2012) and the documentation provided by the enterprise.

1. Create a sense of urgency and make this clear
    - Create a leadership coalition
    - Develop a vision
2. Make it known in the enterprise
    - Communicate the changes on InsideOut
    - Hold information sessions
    - Enlist volunteers
3. Make it real
    - Impact assessment
    - Communicate impact to those affected
    - Perform pilots within project teams
    - Generate small wins
4. Make it happen - deployment
    - Deploy FAIR-iT department per department
    - Provide training and e-learning
    - Accelerate change through larger wins
5. Make it stick - Adapt the culture
    - Provide post-implementation support
    - Create KPI dashboard
    - Celebrate the new organisational culture

The following actions should be undertaken to inform and involve employees. First, a post on InsideOut will be created. Then, the department that will be affected the most will be invited to an information session where the implementation will be explained more in depth. After that, the first pilots will be done to create change ambassadors and check the requirements. Then, the implementation will be performed department per department and during that process, training and guides will be available so that employees can develop the necessary skills for the new system so that they are capable to work with it. Once the implementation is done, an enterprise-wide KPI dashboard will be provided.

The created change management steps, principles, and data preparation flow will be validated through interviews with experts from the enterprise.

Chapter 4

Validation

This chapter summarizes the interviews that were held as part of the validation for the created framework, as well as feedback received from one of the case study participants. The adjustments that will be made to the framework are explained here but can be read in chapter 5. Five experts working in the field of IT consultancy have been interviewed during this phase of the research. The interviews are summarized below.

The participants for the validation of the framework were selected based on availability and years at the enterprise. As can be seen in table 4.1, their experience ranges between half a year to 5 years. The reason why there are no senior consultants with much more experience than 5 years is because of the way the enterprise has structured roles. Most employees will move towards manager after a certain amount of time with the enterprise.

Table 4.1: Validation participants

VP	Position	Experience at their department	Expertise
VP 1	Consultant	2 years	End-to-end projects and financial data
VP 2	Senior consultant	5 years	Implementation projects
VP 3	Senior consultant	3 years	Implementations in European systems
VP 4	Consultant	6 months	No finished projects yet
VP 5	Consultant	1 year	PMO for large implementations

4.1 FAIR-iT principles

VP 1 agreed that internal is an important factor to add to the FAIR principles in order to make them a better fit for the enterprise at hand. They mentioned that because they often work with sensitive client data, it is of crucial importance that this data stays within the enterprise and is not directly shared with competitors. As they have not worked with reusing data a lot, it was difficult for them to say something about whether data quality was important. However, they did note that it would be important to update and maintain the data that is present because their clients'

organisations change frequently in terms of the organisational structure of systems. They wonder if the data itself will ever be reused, but underline that archiving and searching will become a lot quicker.

The opposite of VP 1, VP 2 thinks that trustable is the best addition to the original FAIR principles because it fits well with the core values that the enterprise has identified as part of their organisational culture. For internal, they wonder if the enterprise wouldn't first have to shift their systems to a more cloud-based environment so that data can actually be shared across departments and daughter firms. They did highlight that the internal principle is crucial as the enterprise works with confidential client data and has a large responsibility to ensure the privacy of that data. They suggest the use of an enterprise-wide private cloud to support the implementation of the FAIR-iT principles. They mentioned that these principles will enable the enterprise to become more future-proof as they believe that data sharing will become the norm. Additionally, they mention that the notion of these principles can be incorporated into the annual data training that employees have to complete in order to create more awareness.

VP 3 immediately agrees that trustable is important because it fits the enterprise's core values, as well as the method the department has developed for implementation projects, which has a large focus on connectedness and trust. They point out that every project always has an audit side, and that trustable data is key to performing that task well. They highlight that the department works with private information and that sharing of that data is not desired. They highlight that the trustable principle mostly fits well with the enterprise they work at and that the internal principle is more industry-wide as there currently still exists a lot of competition. They highlight that sharing your data is not desirable at this moment for two reasons: there are strict privacy regulations within the contracts and it might mean giving away your competitive advantage. They share concerns about necessary third-party access and how that would be orchestrated in this system, but after an explanation about licenses, they agree that the risks are minimised in that way. Like VP 2, they agree that there should be an eye for compliance with the GDPR and that the system should have some kind of notification setting that aids the enterprise in remembering this.

VP 4 also highlighted the importance of clearly stating the data the framework is created for. This was confusing to them. As they have little experience with working with client data, they find it hard to say what the direct benefits of the implementation of FAIR-iT would be. They do, however, think that it will be an improvement from the current system as it is not clear and difficult to find data back. Like other VPs, they think that compliance with the GDPR should be clearly stated and that the system should provide some kind of notification once the moment has been reached to delete certain files.

VP 5 mentions that they believe that Internal and trustable are two strong additions. Generally, they believe that the FAIR principles can greatly benefit the enterprise as the current process of data storage is not convenient. They do highlight that clear access rules are crucial, because not all files should be visible to every

employee in the enterprise.

4.2 Data preparation

The created process flow was clear to VP 1. They did, however, mention that a few suggestions on how the steps could be performed would be useful for new employees starting in this job. The explanation that is currently provided is clear, but a few very concrete suggestions would improve clarity. They explained that they would prefer this in the form of a checklist. They did remark that the process was logical to them and applicable to the current flow of data. Additionally, they remarked that the process will become more efficient if the responsibility is shared with more positions than just management. As certain tasks are already shifting roles within their department, it would be logical that the responsibility is also partially shifted to, for example, the PMO.

VP 2 mentions that the process flow is clear but, like VP 1, they would prefer some concrete suggestions. Additionally, they think that the flowchart should mention the data types more specifically, as they wondered whether it was applicable to employee data or client data. They agreed that when the implementation starts, it will not be done retroactively as the workload would be too large. They expressed a desire for a clear example of metadata so that the PMO can confirm that they performed the step correctly. They added the suggestion that Step 2 of the data process needs to incorporate the cleansing of the received data as well so that the data quality increases at the start. This participant, too, expressed the importance of version control and mention that this is currently not organised well enough. Finally, the participant raised concerns about compliance with the GDPR. They mention that it is important that the data archiving and storage is in line with the GDPR, for example, that the data is deleted once the GDPR mentions it should be deleted. This will be something extra that the responsible employee will have to keep track of.

When discussing the data preparation flow, VP 3 remarked that the flow is clear and easy to understand for them. Interestingly, they do note that not every project has a PMO if the engagement is too small. In this case, they believe that the responsibility for the data should remain with the engagement manager. Additionally, they think that a checklist would not be beneficial as every engagement is different from the last, but do agree with other VPs that for some steps, some guidance would be appropriate. An example of this is the creation of metadata, which is something they are not familiar with. They also underline the importance of version control and recognise that metadata can actively contribute to this. They do mention that time is of the essence within their department and that people will not accept a new method if it takes too long before clear benefits are shown. Additionally, they believe that not all data should be stored mindlessly without consideration if it will be needed at a later point. They suggest adding a step to the data preparation flow where at step 2, there is an option to discontinue the process if the file at hand is useless.

The workflow is understandable, but VP 4 would prefer there to be more guidance on how to perform specific steps. They do wonder if the strategy of making the

data responsibility part of the PMO role is a good choice because the PMO is often a consultant who does not have extensive knowledge of data regulations yet. They think that if a consultant takes the PMO role, then a manager should be there to accompany them in the responsibility.

The first remark that VP 5 makes, is that they agree that the PMO role is extremely suitable for this responsibility. Other than VP 4, they believe that this role should be able to handle this responsibility well. As data storage takes a large amount of time, the participant also notes that the solution will be cost-effective due to differences in pay between engagement managers and consultants. Additionally, they remark that engagement managers are extremely busy and that taking this task off their hands will benefit the entire project team. They do highlight that the responsibility should be shared and suggest a four-eye principle. The PMO should take the lead but there should be a manager that checks the work at set moments. They think that the process model is easy to understand and clear, but do highlight that many of the current systems will have to be adapted. Especially the step where licenses are provided should be explained clearly. Step 8 is important to them as that is where a large focus on the principle of trustable can be found back. They express a desire for a part of this task to be automated but recognize that this is not possible in the current state of the systems and applications. Finally, they highlight that the process model would be a good reminder for employees to store the data they work with timely, and not wait until an engagement is finished before everything gets stored.

4.3 Change management

VP 1 mentions that the explained steps for change management are clear and seem effective. As information reaches employees from multiple sources, participant 1 believes that the communication is extensive and well thought-out. They appreciate the focus on making employees capable to work with the new system through training and workshops. Participant 1 expects that employees will accept this system relatively easily as it provides a structure for something that has to be done already anyway. In their eyes, it should make the current process easier. Interestingly, they believe that there should only be a set amount of training provided and that employees should figure things out by themselves too. The culture within the enterprise is that change is often pushed top down, so this change management plan is a welcome adaptation in their eyes. However, they did note that there should not be room for employees to ignore the proposed solution. A certain level of force should be applied where necessary. In one final comment, they mentioned that they would prefer to see a one-pager or poster with the whole change plan on it, in order to create a quick overview as it provides a long-term vision.

The first suggestion VP 2 made was to increase the focus on the shift in ownership and responsibility in the change management plan. Making people enthusiastic about the change and making them part of it will be a useful strategy to do this. They wonder if the pilots should only focus on fixing requirements, and add that it would be a good opportunity for "training the trainer", which fits well with the use of project ambassadors. If you make people knowledgeable as early on as possible,

they can help each other. The project ambassadors can then not only be the spokespeople for the project but can also help in supporting their colleagues. Lastly, they think that review sessions every once in a while after the implementation is done will be useful so that the requirements of the system grow with the organisation and so that employees will continue to be heard.

The steps described in the change management plan are clear to VP 3 and they feel as though every part of a change has been considered. They do, once more, highlight that time is scarce and that it is important to communicate to employees how big the added tasks are and how long it is estimated that it will take them to complete their new tasks. They do think that the risk procedures and internal controls have to be taken into account when managing the change, as it is difficult to change that enterprise-wide. Currently, many files have to be archived manually and it will please employees if a part of this process is made easier. So, they suggest that the focus of the change should be on employee benefits. As a final note, this participant also suggested review sessions after the implementation is done.

VP 4 thinks that the change management plan is extensive and considerate. They do mention that it might be too considerate even and that a certain level of top-down pushing is required to force employees to adapt their way of working. They suggest making the information sessions and training mandatory because otherwise, people will not join. They believe that considerate top-down influence is most effective within the enterprise, even if you have a large ambassador network. They do agree that the pilots and workshops are useful but mention that they should be held earlier in the process to involve employees more in solution design. They suggest that ADKAR is added to the change management plan as it suggests clear steps.

Like VP 3, VP 5 expresses worries related to the amount of time people have available to adjust their way of working. They highlight the importance of the project ambassador remaining available after the implementation is done as a point of contact. They agree that a pilot is important and that it will save a lot of convincing if the implementation itself proves to be beneficial through the leveraging of small wins. They agree that a certain level of force should be applied, and jokingly remarks that they should follow the advice they always give to their clients: accept the change. They agree that the change management plan provides a long-term vision. Lastly, they highlight that they think that the shifting of responsibility towards the PMO role provides a valuable opportunity for consultants to show that they are up for the task and able to take such a responsibility. However, they do note that a shift in mindset will be required for managers before they can let go of the task as they are responsible for the engagement overall.

4.4 Case study respondent

One case study participant responded to the request for feedback on the presented principles and change management plan. They mention that they doubt whether principle T2.1 should not be T3 instead. They believe that the quality of metadata is important to such an extent that it should be its own principle. Additionally, they think that updates of metadata should be seen separately from the enter-

prise standards, thereby giving another reason to change the principle to T3. They do, however, remark that frequently updating metadata and providing provenance should be an enterprise standard. Additionally, they agreed with the validation interview participants and underlined that the change management plan should fit their enterprise strategy well. They suggested making a video to communicate the changes to employees rather than a poster.

4.5 Summary and adjustments

Overall, participants agreed that the addition of the FAIR principles will be beneficial to the enterprise they work for. Additionally, they acknowledged that the addition of the "iT" makes them more applicable to the enterprise context. Participants disagreed about whether the PMO role is suitable for the responsibility over the data. Participant 4 expresses worries that the employees taking this role will have too little experience, whereas Participant 5 thinks it will be a great opportunity for employees to prove that they can handle that type of responsibility. Most participants agreed that the PMO role was the best position to add responsibility to. They do agree that a four-eye principle would be beneficial, as the responsibility is large and the consequences if the tasks are not performed correctly are severe.

Some validation participants requested specific guidelines for data preparation. As it is a framework that is being developed and not a procedure, there will be no step-by-step guide on what steps to exactly undertake to prepare the data. As Participant 3 remarked: not every project is the same so a task list would not necessarily be beneficial. However, certain suggestions for the steps can definitely be added to guide employees in their tasks. Participant 5 expressed technical worries that fall out of the scope of this thesis, but they will be taken into account in the future work section in section 6.3.3. Participant Three made the striking remark that not every file is as important as others, thereby suggesting the discontinuation of the process in the case that a file is useless. They suggest focusing that the change management focuses more on reaping benefits to create enthusiasm, which is something that the project ambassadors can provide. GDPR compliance is very important to these people and should be taken into account.

The request for a poster of the change management plan is a good remark, as this provides clarity to employees as well as the long-term vision that one case study participant suggested. The suggestion of adding ADKAR to the change management plan will be taken into account as this provides an angle that Kotter and the enterprise model do not. Additionally, it is mentioned the training of trainers as part of project ambassadors. Their last remark on the review sessions is important as well because working with a system officially starts once the change is done. Adding this to the change management plan will stretch it out into actually working with the system. These suggestions will be taken into account for the final change management plan.

Relating the aforementioned validations back to Wieringa (2014), most of the factors of the treatment validation have been satisfied. The factors are: the artefact produces the desired effects, what are the trade offs for other artefacts, and do the

effects satisfy the requirements The artefact, according to VPs, produces the wanted effect while satisfying requirements. Additionally, trade-offs have been made for the different artefacts, such as the data preparation flow that will be altered based on feedback from participants.

Chapter 5

FAIR-iT framework

This chapter contains the final version of the framework. As validation participants agreed that the principles are accurate and advantageous, they will not be changed. Therefore, only a short overview of the six principles will be given. For the explanation of the principles, please see section 3.2. The adaptions of the data preparation and change management components will be discussed here. The new data preparation model will be shown and explained through suggestions for actions as requested by the validation participants. Additionally, the change management steps have been adapted and will be shown in a one-pager for long-term focus and clarity.

5.1 The FAIR-iT principles

In figure 5.1, the six FAIR-iT principles can be found. Important to note once more is that the "internal" principle is lowercase as case study and validation participants have expressed the belief that there will be a time when the IT sector will cooperate rather than compete. The other principles will then remain the same, but the data will not be purely internal anymore.

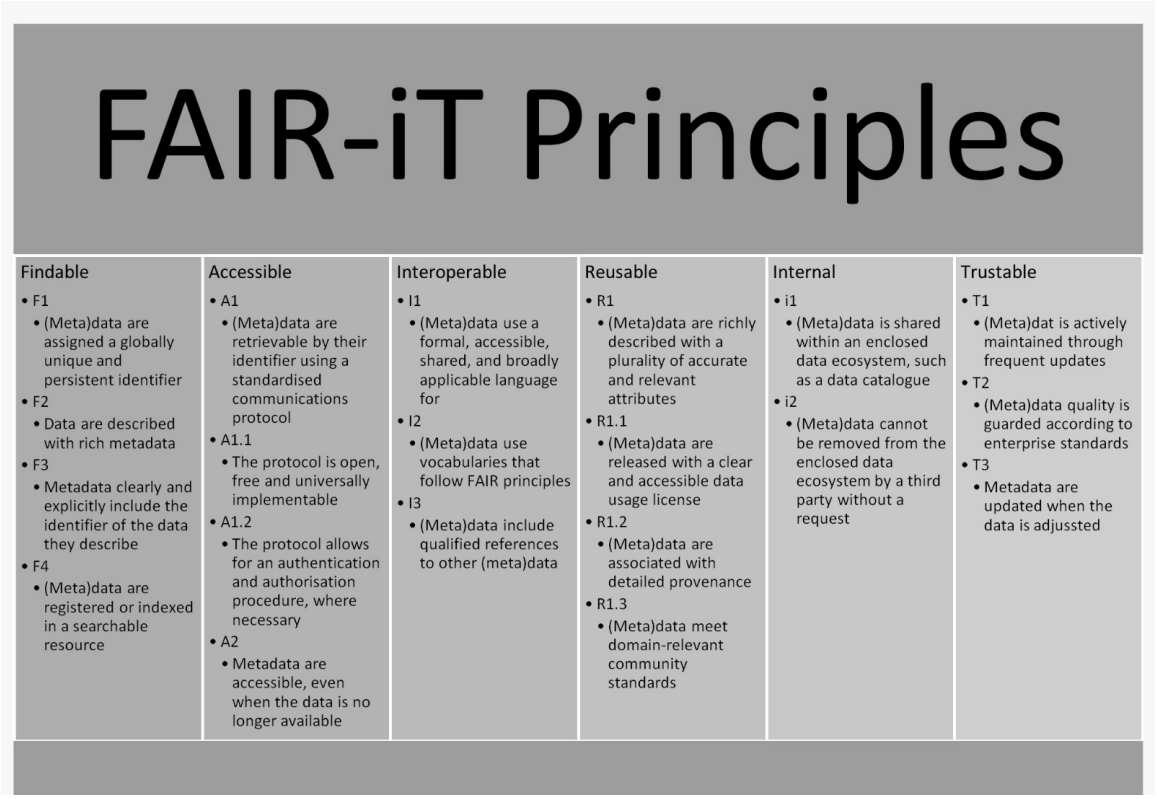

Figure 5.1: The FAIR-iT principles

5.2 The data preparation flow

One participant remarked that the process flow should have a step where it is decided whether the file at hand is useful for saving in a FAIR-iT way. This comment has been taken into account and the new process flow can be seen in appendix I. As can be seen in figure I.1, step two can now flow into step 3, or in step 2a where the process gets terminated. Most of the validation participants suggested that a checklist with steps to undertake in the preparation of data would be beneficial to them. As the artefact that is developed is a framework, no concrete steps will be provided as that limits the applicability of the framework. However, some suggestions per step will be provided to guide employees depending on the task at hand, taken from Jacobsen, Kaliyaperumal, et al. (2020). Please note that these suggestions are all based on the assumption that the actions are within the contractual agreement. Employees should not perform actions that break contract rules.

Table 5.1: Suggestions per step of the data preparation flow

Step	Suggestions
Step 1: Retrieve data from the source	Download files from email Extract files from client system Assess what data is available
Step 2: Analyse data	Assess data representation Assess contents of the data Assess whether the data is useful for the project Check if the data already contains FAIR-iT features like identifiers
Step 3: Research and analyse metadata	Investigate if there is metadata Assess whether metadata describes the right data Identifiy what metadata should be gathered Assess if metadata already contains FAIR-iT features like identifiers
Step 4: Define the semantic model for data & metadata	create a conceptual model Define the ontology Define the class structure and properties Create the semantic model
Step 5: Make the (meta)data linkable	Add appropriate identifiers Use the Resource Description Framework Use a FAIRifier
Step 6: Assign an appropriate license	Define contract terms Classify contract with relation to access rules Define authorised accounts
Step 7: Make data available	Upload data to the database
Step 8: Update & maintain data	Do weekly check-ins during a project Set up notifications for removing data to comply with GDPR When a new version is available, upload it immediately

5.3 The change management guide

As suggested by one validation expert, the ADKAR model for change was added to the change management guide. The titles of the steps have been adapted to better fit this model. The choice to add this model was because it provides more concrete implications for employees while the frameworks are more high-level. As a change management plan is written to guide an organisation, the employees should not be

forgotten. For an explanation of the ADKAR model, see appendix C. The one-pager with the improved steps can be seen in Appendix J. Compared to the steps written in the draft change management guide, certain terms have been adapted and added. There is an increased focus on employee involvement and the opportunity to review the implementation at step 5.

Chapter 6

Discussion

6.1 Implications

The objective of the thesis was to research whether enterprises could benefit from the use of the FAIR guiding principles. This led to the creation of a framework that guides enterprises in their journey towards the adoption of the FAIR principles. In order to do so, requirements have been identified that are applicable to the specific enterprise the thesis was performed at. One requirement was that certain aspects of the FAIR principles have to be highlighted to move the focus point to data quality and protection of data as that aids in making the FAIR principles more applicable for an enterprise context according to case study participants. The encountered evidence has several implications for the enterprise and future research. For the validation of the research, please see Appendix K.

6.1.1 Adoption of FAIR-iT in enterprises

Experts from the case study and validation interviews have expressed interest in learning more about the principles and what they can mean to their departments. The final artefact that was designed has the goal of helping the enterprise at hand prepare its employees for a new corporate culture, which includes a large shift in responsibility over data assets. Looking at the results of the research, the real question enterprises should ask themselves is not how they can implement systems for data storage, but rather how they can adapt these systems to their needs and organisational culture so that the maximum amount of benefits can be experienced. During the research, it has become clear that there are many factors that play a large role in the implementation of the FAIR principles. It is more than only implementing the technical changes and sending an email to the employees. The corporate culture, the organisational structure and the current mindset and skills of employees might be just as important and are therefore the focus point of the framework. As mentioned in the validation interviews, other enterprises might require other principles that fit well with their core values and corporate culture. As mentioned earlier, the shift in mindset with regard to data is likely to be a challenging task. As long as data is seen as power for the individual, rather than power for the organisation, nothing will change. This is why the use of a change management plan with a focus on employee mindset is relevant to the framework. Changes and effects made to the data flow and the addition of a framework are measurable and tangible, for instance in an increase in productivity. Change made to corporate culture, however, is not

as factual and difficult to quantify as this is not easy to measure. Employees might need time to adopt the new corporate culture and make it their own, or it might be in stride with their personal beliefs. The differences between these two aspects of incorporating the FAIR-iT principles in an enterprise are important to keep in mind as they will have to be managed differently.

6.1.2 The future of FAIR in the enterprise context

As identified by case study participants and validation interview experts, the use of the FAIR principles has benefits in an enterprise context. Case study and validation interview participants have highlighted several times that the current state of the industry and the way they work with competitors make data sharing difficult. This makes it challenging to use the FAIR principles to their full extent, as the focus will have to lay on internal data sharing rather than external data sharing at first. The notion of competitiveness within the sector is an important factor as to why the original FAIR principles had to be adapted by adding another letter. It is likely that the proposed process will change in the future as more enterprises embrace the use of artificial intelligence and start to automate tasks through increased machine-readability. With data becoming more and more powerful, it is imperative that enterprises take control of their data and utilise this to its full extent. It is important that employees understand the data they work with and are capable to manipulate the data in beneficial ways. This might require the current employees to up their technical skills, or perhaps hire new employees that already possess this knowledge. It is important to note that the statements made are based on a case study and validation performed at one company. It might be possible that the context in another enterprise differs largely, which means that there will be different implications.

6.2 Limitations of the study

One important limitation is one that has been mentioned several times before. A significant lack of relevant literature on the topic of FAIR in an enterprise context exists. This leads to some parts of this thesis being not as well supported by literature as preferred, showing a lack of references for some parts of the literature review. Additionally, as many companies actively try to be secretive about their assets, little is known about how businesses manage their data. To the best of the author's knowledge, there is not much written in the academic domain about the use of these principles in consultancy enterprises. However, that does not mean that there are no companies using the principles.

A large limitation is the notion that the framework might only be applicable to the enterprise that the thesis was conducted at. Additionally, as it only provides guidance on the preparation of adoption, little technical detail and specific actions are provided. While this is exploratory research in a field that is still in its infancy, the ultimately desired result has not been achieved. Preferably, the thesis would have consisted of an entire implementation program, but this was not possible due to a variety of constraints and expectation management from the author. In order to verify whether the framework brings the expected results, it should be tested

through a case study where the framework is tested in the real world, meaning that a naturalistic evaluation still has to be performed.

The third limitation is the use of subjective measurements in the case study and validation interviews. Measures were taken to limit the subjectivity, but the data collection process was still subjected to it.

6.3 Contributions to Research and Practice

6.3.1 Contributions to research

During the literature review performed in chapter 2, it became clear that there is very little literature available that relates the FAIR principles to the business sector, let alone IT consultancy and audit. This makes the creation of this framework relevant, as it presents a new topic and new opportunities for further development. However, the significant lack of literature also indicates that it might be challenging to implement something in the business industry that is starting to be used frequently in research, as the stakes for both groups can be deemed to be different, as became clear during the case studies. Due to the empirical component, the study provides new insights in the possibilities for the FAIR principles in an enterprise context. Hopefully, this research contributes significantly to the FAIR community and can be used as the foundation for more research on the use of the FAIR principles in other industries than research.

6.3.2 Contributions to practice

This thesis provides a framework based on empirical evidence that can guide project teams in the preparation for the adoption of the FAIR guiding principles. The benefits of the framework have been acknowledge by case study participants and experts who validated the framework, which sparks the hope that the framework will provide them with useful information.

6.3.3 Future work

As mentioned in chapter 3, the framework is only applicable to the preparation for adoption. A method for the implementation of the FAIR-iT principles is still required. The reason why this is not part of this thesis, is the lack of technical knowledge on system integration from the author. Additionally, not all enterprises are ready on a technical level to implement these principles, for instance because there is no data catalogue or because no steps towards automation have been taken thus far. In this case, a large portion of technical preparation is still required.

Additionally, the developed FAIR-iT framework is yet to be tested in other enterprises than the one the thesis was performed at. As not every enterprise is the same on a cultural and technical level, adaptations could be required to make the framework more broadly applicable. Furthermore, tests should be performed in a real-world case study to further define requirements and give guidance on the adoption of the FAIR principles.

Lastly, a maturity model could be developed that aids enterprises in determining the benefits they want to achieve and their current position in relation to the FAIR principles. This could aid in choosing the direction of future growth and help create metrics and KPIs, thereby creating a clearer path for the adoption of the FAIR(-iT) principles.

Chapter 7

Conclusion

7.1 Research questions

This study aimed at researching possibilities for companies to incorporate the FAIR guiding principles into the process of storing engagement files without disrupting the current workflow too much. Secondly, it researched how the integration of FAIR could be managed on a change level to ensure employee resistance was kept at a minimum. The thesis followed the following main question:

What adaptations must be made to the FAIR data principles so that they are applicable to an enterprise context, specifically so that a framework can be developed that allows organisations to prepare the adoption of FAIR into the data storage process of past engagement files in IT consultancy and audit departments?

To answer the main research question, a structured literature review and 4 case studies have been conducted. The results of the literature review and case studies were put into a draft framework, which was evaluated by 5 experts and one case study respondent. The main research question was split up into three sub-questions.

1. *What challenges should the implementation of FAIR solve and how is this currently orchestrated in an enterprise context?*

The structured literature review has shown that the implementation of FAIR can aid organisations in managing their data and governing the assets they possess. According to the literature, enterprises can have difficulty managing their data assets as they are growing continuously, while the skills of their employees might not grow at the same rate. Case study respondents pointed out that their enterprise encounters difficulties in the strategy of saving data from engagements. There is no clear structure on how to do this and this is where FAIR could improve their workflow. The incorporation of FAIR into the storage process of engagement data could make the process more structured and will incorporate more employees into the data storage process, thereby creating more data citizens. This helps enterprises create a shared responsibility over the data they must safeguard. Additionally, the implementation of the FAIR guiding principles should aid in the challenge of making the data more findable and reusable once it has been stored. Case study participants remarked that the current inventory system makes it near impossible to find important files back. It was remarked by case study participants that the FAIR principles did not

accentuate aspects that are crucial for enterprises, such as data quality. This is why the choice was made to add two principles that are not on their own but rather highlight aspects that can be found in the four foundational principles of FAIR.

2. *What frameworks and tools have been developed to aid the adoption and implementation of the FAIR principles and thereby help organisations go FAIR?*

The structured literature review has shown that there are, to the best of the author's knowledge, several options to support organisations with their FAIR data. However, there is not one clear method to follow when implementing the FAIR guiding principles. GO-FAIR proposed the use of workshops and provide a direction for organisations to take when adapting to the FAIR principles, but there is no checklist to follow. There are several tools and workflows that explain how to make FAIR data, which is something of importance that was highlighted by the case study and validation participants. A data preparation flow was proposed to help employees in their shift towards data responsibility and to support them in their objective to create FAIR-T data.

3. *What change management steps should enterprises undertake to guide the adoption and implementation of the FAIR principles?*

The final section of the created framework answered the final sub-question. A change management plan was proposed, supported by a framework that was provided by the enterprise the research was performed at as well as several frameworks found in the literature. The keynote in creating the change management plan was to not forget the employees and to take them along in the journey from the start. This can have many applications. Literature and several case study and validation interview participants highlighted the importance of workshops and training, whereas others remarked that providing a long-term vision is of significant importance in such a plan. The result from this sub-question is a one-pager with change management steps, which can be found in Appendix J.

For this thesis, a combination of several research methods was used. A structured literature review was performed, as could be read above. Additionally, a case study at one company was held, which was validated through interviews with different employees at the same enterprise. Interestingly, most participants gave highly similar answers to questions about their encountered challenges and ideas for a solution. There was not a single participant that mentioned that they thought that the FAIR principles would not solve (at least a part of) their challenges with the process of data storage and archiving. Importantly, most participants did mention concerns about the implementation on a technical level.

7.2 Key contributions

During the structured literature review, a gap in empirical research on the use of FAIR in an enterprise (and specifically IT consultancy and audit) context was identified. The study at hand aims to help researchers and practitioners realise the benefit for enterprises that is currently still hidden in the FAIR guiding principles. The findings in this research align with the limited number of available resources. The main contributions to research and practice are:

1. *Academic:* This study provides empirical research on a topic that is still in its infancy: the use of the FAIR principles outside of an academic context. As the used literature has shown, the FAIR principles thus far have not been extensively researched outside of an academic context. This thesis provides the basis for other research to elaborate on and will hopefully grow into a blooming branch of research, where academics have the potential to focus on the implementation and implications of these principles in a new context. Hopefully, the framework and ideas presented in this thesis will inspire other researchers to place the FAIR principles in new contexts, outside of research, so that a large knowledge base of opportunities can come into being.

2. *Practical:* This study provides a framework that enterprises can employ in their process of becoming FAIR. It provides a model that employees can benefit from in their daily tasks. Additionally, empirical research has shown that there is interest within enterprises to change processes and prepare for (partial) automation of tasks related to engagement data storage.

3. *Practical:* This study provides insights into the wishes of (senior) managers of IT audit and consultancy departments concerning the developments in their professional sector. Enterprises are more secretive about their processes and data to maintain a competitive advantage. This is why the FAIR principles will be useful for internal use at first. However, as several case study participants noted: there might be a future where there is room for collaboration rather than competition, and that is where the use of the FAIR-iT principles will flourish; once they become the FAIR-T principles. It might even serve as a call to enterprises to enable themselves to cooperate rather than compete, as one case study participant remarked: "We are stronger together".

References

Akrong, G. B., Shao, Y., & Owusu, E. (2022). Overcoming the challenges of enterprise resource planning (erp): A systematic review approach. *International journal of enterprise information systems (IJEIS)*, *18*(1), 1-41. doi: http://doi.org/10.4018/IJEIS.306242

Alballaa, H., & Al-Mudimigh, A. (2011). Change management strategies for effective enterprise resource planning systems: a case study of a saudi company. *International journal of computer applications*, *17*(2), 14-19.

Amoako-Gyampah, K., & Salam, A. (2004). An extension of the technology acceptance model in an erp implementation environment. *Information and Management*, *41*(6), 731-745.

Bharathi, S. V., & Mandal, T. (2015). Prioritising and ranking critical factors for sustainable cloud erp adoption in smes. *Int. J. of Automation and Logistics*, *1*. doi: 10.1504/IJAL.2015.071723

Bharathi, S. V., Vaidya, O., & Parikh, S. (2012). Prioritizing and ranking critical success factors for erp adoption in smes. , *6*.

Braunschweig, K., Eberius, J., Thiele, M., & Lehner, W. (2012). The state of open data - limits of current open data platforms.

Bryman, A. (2012). *Social research methods* (4th ed.). Oxford University Press.

Budroni, P., Claude-Burgelman, J., & Schouppe, M. (2019). Architectures of knowledge: The european open science cloud. *ABI Technik*, *39*(2), 130-141. doi: https://doi.org/10.1515/abitech-2019-2006

Burgelman, J.-C., Pascu, C., Szkuta, K., Von Schomberg, R., Karalopoulos, A., Repanas, K., & Schouppe, M. (2019). Open science, open data, and open scholarship: European policies to make science fit for the twenty-first century. *Frontiers in Big Data*, *2*. Retrieved from https://www.frontiersin.org/articles/10.3389/fdata.2019.00043 doi: 10.3389/fdata.2019.00043

Channon, D. F., & Caldart, A. A. (2015). Mckinsey 7s model. *Wiley encyclopedia of management*, 1-2.

Commission, E., for Research, D.-G., & Innovation. (2016). *Realising the european open science cloud : first report and recommendations of the commission high level expert group on the european open science cloud.* Publications Office. doi: doi/10.2777/940154

da Silva Santos, L. O. B., Burger, K., Kaliyaperumal, R., & Wilkinson, M. D. (2023, 03). FAIR Data Point: A FAIR-Oriented Approach for Metadata Publication. *Data Intelligence*, *5*(1), 163-183. Retrieved from https://doi.org/10.1162/dint_a_00160 doi: 10.1162/dint_a_00160

Farrier, E. (2022). What is metadata: Definition and meaning. *Avast*.

Garcia, L., Batut, B., Burke, M. L., Kuzak, M., Psomopoulos, F., Arcila, R., ... Palagi, P. M. (2020). Ten simple rules for making training materials fair. *PLoS Computational Biology*, *16*(5). Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-85085156809&doi=10.1371%2fjournal.pcbi.1007854&partnerID=40&md5=a457775567a1b12e433a89e12e797da2 doi: 10.1371/journal.pcbi.1007854

Garousi, V., Felderer, M., & Mäntylä, M. V. (2017). Guidelines for conducting multivocal literature reviews in software engineering.

George, G., Haas, M. R., & Pentland, A. (2014). Big data and management. *Acad. Manage. J.*, *57*(2). doi: 10.5465/amj.2014.4002.

Glaser, B., & Strauss, A. (1967). The discovery of grounded theory: Strategies for qualitative research.

GO-FAIR. (2018). Fair principles. Retrieved from https://www.go-fair.org/fair-principles/

Gryk, M. R., Schuyler, A., Wedell, J., Maciejewski, M., Weatherby, G., & Romero, P. (2019). When should fair begin? In *2019 acm/ieee joint conference on digital libraries (jcdl)* (p. 426-427). doi: 10.1109/JCDL.2019.00102

Gualtieri, M. (2016). The forrester wave: Big data hadoop distributions.

Hauschke, C., Nazarovets, S., Altemeier, F., & Kaliuzhna, N. (2021). Roadmap to fair research information in open infrastructures. *Journal of Library Metadata*, *21*(1-2), 45 – 61. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-85118461644&doi=10.1080%2f19386389.2021.1999156&partnerID=40&md5=d2847b970a7ea71b8ee1405d5a75cce8 doi: 10.1080/19386389.2021.1999156

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in is research. *MIS Quarterly*, *28*(1), 75–105.

Hiatt, J. (2006). Adkar: a model for change in business, government, and our community.

Inau, E. T., Sack, J., Waltemath, D., & Zeleke, A. A. (2021). Initiatives, concepts, and implementation practices of fair (findable, accessible, interoperable, and reusable) data principles in health data stewardship practice: Protocol for a scoping review. *JMIR Research Protocols*, *10*(2). Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-85100459804&doi=10.2196%2f22505&partnerID=40&md5=4a21d3d0b928fdf2477771317a49a895 doi: 10.2196/22505

Jacobsen, A., de Miranda Azevedo, R., Juty, N., Batista, D., Coles, S., Cornet, R., & Schultes., . . E. (2020). Fair principles: Interpretations and implementation considerations. *Data Intelligence*, *2*, 10–29. doi: 10.1162/dint_r_00024

Jacobsen, A., Kaliyaperumal, R., Santos, L. O. B. d. S., Mons, B., Schultes, E., Roos, M., & Thompson, M. (2020). A generic workflow for the data fairification process. *Data Intelligence*, *2*(1-2), 56 – 65. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-85117835056&doi=10.1162%2fdint_a_00028&partnerID=40&md5=d21ed421a118c21af968b5fc61210b46 doi: 10.1162/dint_a_00028

Juty, N., Wimalaratne, S. M., Soiland-Reyes, S., Kunze, J., Goble, C. A., & Clark, T. (2020). Unique, persistent, resolvable: Identifiers as the foundation of fair. *Data Intelligence*, *2*(1-2), 30 – 39. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-85089273603&doi=10.1162%2fdint_a_00025&partnerID=40&md5=6f6ffc30c15c44ef718b588e9a3b741e doi: 10.1162/dint_a_00025

Katz, D. S., Psomopoulos, F. E., & Castro, L. J. (2021). *Working Towards Understanding the Role of FAIR for Machine Learning*. Zenodo. Retrieved from https://doi.org/10.5281/zenodo.5594990 doi: 10.5281/zenodo.5594990

Kerssens, M. (2023). Research topics: Ethical challenges encountered in the use of big data analytics in erp implementation - final version.

Kitchenham, B. (2004). Procedures for performing systematic reviews. *Software Engineering Group Keele University Empirical Software Engineering*.

Kotter. (2012). The 8-step process for leading change. Retrieved from https://www.kotterinc.com/8-step-process-for-leading-change

Kotter, J. P. (1995). Leading change: Why transformation efforts fail. *Harvard Business Review*, *73*(2), 59–67.

Kotter, J. P. (1996). Leading change.

Krechmer, K. (2005). The meaning of open standards. In *Proceedings of the 38th annual hawaii international conference on system sciences* (p. 204b-204b). doi: 10.1109/HICSS.2005.605

Krechmer, K. (2008). Open standards: A call for action. In *2008 first itu-t kaleidoscope academic conference - innovations in ngn: Future network and services* (p. 15-22). doi: 10.1109/KINGN.2008.4542244

Labadie, C., Legner, C., Eurich, M., & Fadler, M. (2020). Fair enough? enhancing the usage of enterprise data with data catalogs. In *2020 ieee 22nd conference on business informatics (cbi)* (Vol. 1, p. 201-210). doi: 10.1109/CBI49978.2020.00029

Landi, A., Thompson, M., Giannuzzi, V., Bonifazi, F., Labastida, I., da Silva Santos, L. B., & Roos, M. (2020). The "a" of fair – as open as possible, as closed as necessary. *Data Intelligence*, *2*, 47–55. doi: 10.1162/dint_a_00027

Lin, D., Crabtree, J., Dillo, I., Downs, R. R., Edmunds, R., Giaretta, D., ... Westbrook, J. (2020). The trust principles for digital repositories [journal]. *Scientific Data*. Retrieved from https://doi.org/10.1038/s41597-020-0486-7 doi: 10.1038/s41597-020-0486-7

Masheshwar, C., & Javalagi, C. (2019). An assessment of factors affecting erp implementation: A systematic view. *International Journal of recent technology and engineering*, *8*(3), 3801-3808.

Mesicek, L., Petrus, P., & Stranska, A. (2021). Expected and achieved benefits of erp systems implementation over time. *Acta Informatica Pragensia*, *10*(1), 22-37. doi: 10.18267/j.aip.144

Mons, B., Neylon, C., Velterop, J., Dumontier, M., da Silva Santos, L. O. B., & Wilkinson, M. D. (2017). Cloudy, increasingly fair; revisiting the fair data guiding principles for the european open science cloud. *IOS Press*, *37*, 49-56. doi: 10.3233/ISU-170824

Muilenburg, J. (2021). Persistent identifiers - about persistent identifiers. *University of Washington*.

Murray-Rust, P. (2008). Open data in science. *Nature Precedings*. Retrieved from https://doi.org/10.1038/npre.2008.1526.1 doi: 10.1038/npre.2008.1526.1

Musen, M. (2019). Making data fair requires more than just principles: We need knowledge technologies. In *2019 15th international conference on escience (escience)* (p. 530-532). doi: 10.1109/eScience.2019.00071

Nicholson, C., Kansa, S., Gupta, N., & Fernandez, R. (2023). Will it ever be fair?: Making archaeological data findable, accessible, interoperable, and reusable. *Advances in Archaeological Practice*, *11*(1), 63–75. doi: 10.1017/aap.2022.40

O'hara, K. (2019, February). *Data trusts: Ethics, architecture and governance for trustworthy data stewardship* (Project Report No. 10.5258/SOTON/WSI-WP001). Retrieved from https://eprints.soton.ac.uk/428276/

Pană, G. T., Ivanoaica, T., Raportaru, M. C., Băran, V., & Nicolin, A. (2021). Towards the implementation of fair principles on an earthquake analysis platform. In *2021 20th roedunet conference: Networking in education and research (roedunet)* (p. 1-4). doi: 10.1109/RoEduNet54112.2021.9638283

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2008). A design science research methodology for information systems research. *Journal of Management Information Systems*, *24*(3), 45–77. doi: doi:10.2753/MIS0742-1222240302.

Plotkin, D. (2020). *Data stewardship: An actionable guide to effective data management and data governance.*

Ren, X., Li, X., Deng, K., Ren, K., Zhou, A., & Song, J. (2020). Bringing semantics to support ocean fair data services with ontologies. In *2020 ieee international conference on services computing (scc)* (p. 30-37). doi: 10.1109/SCC49832.2020.00011

Riege, A. M. (2003). Validity and reliability tests in case study research: A literature review with "hands-on" applications for each research phase. *Qualitative Market Research: An International Journal*, *6*(2), 75–86. doi: 10.1108/13522750310470055.

Robinson, E., Buys, M., Chodacki, J., Garzas, K., Monfort, S., Nancarrow, C., ... Davies, N. (2023, 03). FAIR Island: real-world examples of place-based open science. *GigaScience*, *12*. Retrieved from https://doi.org/10.1093/gigascience/giad004 doi: 10.1093/gigascience/giad004

Rosenbaum, S. (2010). Data governance and stewardship: Designing data stewardship entities and advancing data access. *Health Services Research*, *45*(5p2), 1442-1455. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-6773.2010.01140.x doi: https://doi.org/10.1111/j.1475-6773.2010.01140.x

Roszkiewicz, R. (2010). Enterprise metadata management: How consolidation simplifies control. *Journal of Digital Asset Management*, *6*(5), 291-297. doi: 10.1057/dam.2010.32

Russell, K. (2021). *Fair principles and why they matter*. doi: https://doi.org/10.1002/9783527825042.ch5

Russom, P. (2017). The data catalog's role in the digital enterprise: Enabling new data-driven business and technology best practices. *TDWI Consultancy Report*.

Shah, S., Bokhari, R., Hassan, S., Shah, M., & Ali, M. (2011). Socio-technical factors affecting erp implementation success in pakistan: an empirical study. *Australian journal of basic and applied sciences*, *5*(3), 742-749.

Shanahan, H., & Bezuidenhout, L. (2022). Rethinking the a in fair data: Issues of data access and accessibility in research. *Frontiers in Research Metrics and analytics*, *7*. doi: 10.3389/frma.2022.912456

Supramaniam, M., & Kuppusamy, M. (2011). Analysis of critical success factors in implementing enterprise resource planning systems in malaysian business firms. *THE ELECTRONIC JOURNAL OF INFORMATION SYSTEMS IN DEVELOPING COUNTRIES*, *46*(1), 1-19. doi: https://doi.org/10.1002/j.1681-4835.2011.tb00327.x

Tarhini, A., Ammar, H., Tarhini, T., & Masa'deh, R. (2015). Analysis of the critical success factors for enterprise resource planning implementation from stakeholders' perspective: A systematic review. *International business research*, *8*(4), 25-40. doi: https://doi.org/10.5539/ibr.v8n4p25

Thompson, M., Burger, K., Kaliyaperumal, R., Roos, M., & Santos, L. O. B. d. S. (2020). Making fair easy with fair tools: From creolization to convergence. *DATA INTELLIGENCE*, *2*(1-2, SI), 87-95. doi: 10.1162/dint_a_00031

Tome, L. . N.-M. M., A. K. M. A. (2014). Barrier to open source erp adoption in south africa. *The African Journal of Information Systems*, *6*(2), 26-47.

Van Hau, T., & Kuzic, J. (2010). Change management strategies for the successful implementation of enterprise resource planning systems. *IEEE - Second international conference on knowledge and systems engineering*, 178-182.

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). Feds: a framework for evaluation in design science research. *European Journal of Information Systems*, *25*(1), 77-89. Retrieved from https://doi.org/10.1057/ejis.2014.36 doi: 10.1057/ejis.2014.36

Waterman, R. (1982). The seven elements of strategic fit. *Journal of Business Strategy*, *3*, 68–72.

Waterman, R., Peters, T., & Phillips, J. (1980). Structure is not organization. *Business Horizons*, *23*(3), 14–26.

Wieringa, R. J. (2014). Design science methodology for information systems and software engineering.
doi: doi:10.1007/978-3-662-43839-8

Wilkinson, M., Dumontier, M., & Aalbersberg, e. a., I. (2016). The fair guiding principles for scientific data management and stewardship. *Sci Data*, *3*. doi: https://doi.org/10.1038/sdata.2016.18

Wilkinson, S. R., Eisenhauer, G., Kapadia, A. J., Knight, K., Logan, J., Widener, P., & Wolf, M. (2022). F*** workflows: when parts of fair are missing. In *2022 ieee 18th international conference on e-science (e-science)* (p. 507-512). doi: 10.1109/eScience55777.2022.00090

Wise, J., de Barron, A. G., Splendiani, A., Balali-Mood, B., Vasant, D., Little, E., ... Hedley, V. (2019). Implementation and relevance of fair data principles in biopharmaceutical rd. *Drug Discovery Today*, *24*(4), 933-938. Retrieved from https://www.sciencedirect.com/science/article/pii/S1359644618303039 doi: https://doi.org/10.1016/j.drudis.2019.01.008

Wixom, B., & Ross, J. (2017). How to monetize your data. *MIT SloanManag. Rev.*, *58*(3).

Wolf, M., Logan, J., Mehta, K., Jacobson, D., Cashman, M., Walker, A. M., ... Cliff, A. (2021). Reusability first: Toward fair workflows. In *2021 ieee international conference on cluster computing (cluster)* (p. 444-455). doi: 10.1109/Cluster48925.2021.00053

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE.

Appendix A

Structured Literature review: Search Protocol

A.1 Inclusion and Exclusion criteria academic search

Inclusion	Exclusion
Papers about FAIR frameworks	Papers only about FAIR benefits
Papers about implementing FAIR	Papers only about FAIR disadvantages
Papers about FAIR data management	Papers with less than 15 citations
Papers about FAIR in IT consultancy	

A.2 Search results

Databases used: Scopus, Web of Science, IEEE Xplore, ACM Digital Library, and ScienceDirect.
Search key: *"FAIR principles"* AND (*framework* OR *implementation* OR *management* OR *"IT consultancy"*).
Search performed on 07.06.2023.

Search process	Amount of papers
Unique Results	325
After filtering on title	53
After filtering on abstract	21
After filtering on full text	14
Adding forward & backward references	18

Appendix B

Structured Literature review: Results

Table B.1 shows the papers that are selected for the structured literature review. A short description of the paper is given, as well as the title and the type of research. Lastly, the table also contains the types of papers. There are 6 journal papers (J), 7 conference papers (C), 1 article (A), and one white paper (W).

B.1 Selected papers

Table B.1: The selection of papers from the systematic review

Paper	Type	Description
Roszkiewics (2010)	J	Tips for implementation and management of metadata and data catalogues. No clear conclusion based on validated research.
Labadie et. Al. (2020)	C	Enterprise data must become FAIR as it is a strategic asset. Data catalogues are a way to realise this. They are, however in their infancy and not well-defined. A qualitative study has been performed on data catalogue initiatives. Findings contribute to FAIR research in enterprises.
Ren et. Al. (2020)	C	FAIR principles are beneficial to ocean data management. Data is currently not produced and managed well. Novel techniques are proposed to solve this through the use of a unified semantic model. This is optimised with the use of data structures.
Wilkinson et. Al. (2022)	C	The FAIR principles are becoming widely used but the level of FAIRness is difficult to define. Additionally, there is too much focus on findable but other principles are lacking. New tools are required. A workflow is proposed.

Musen et. Al. (2019)	C	FAIR led to improvements in data storage. However, the principles are not enough. New knowledge technologies are required to aid in making data FAIR.
Wolf et. Al. (2021)	C	Current workflows are not necessarily FAIR. Technical debt is recognized as a key factor in implementing new workflows and systems. Performed several experiments to test their model, among which were simulations.
Pana et. Al. (2021)	C	They present a FAIR approach for an analysis platform. The tools and process for FAIRification of seismic data are proposed. They propose FAIRification dependent on data ingestion pipelines, which satisfies the set requirements.
Gryk et. Al. (2019)	C	Research data communities recognize FAIR data principles as the standard. While they are useful, their applicability is unclear in some cases. Research is performed on the value of FAIR at the point of creation, rather than at storage.
Jacobsen et. Al. (2020)	J	Interpretation and implementation considerations are provided. These should assist in accelerated global participation in FAIR.
Jacobsen et. Al. (2020)	J	A FAIRification workflow is proposed and it consists of workshops and steps to undertake. Each step is described with processing, required expertise and tools
Hauschke et. Al. (2021)	J	The FAIR principles can be applied to several contexts. It is researched how this can be made easier in an open infrastructure so that more stakeholders can benefit
Garcia et. Al. (2020)	J	FAIR can be applied to training programs. A workflow is proposed on how to do this.
Juty et. Al. (2020)	J	The FAIR principles require identifiers to be findable on the web. These should be associated with metadata. Basic principles for identifier design are proposed.
Thompson et. Al. (2021)	A	Questioning what is FAIR and how organisations can become FAIR. They show examples of tools and workflows under development so that these can be combined at a later point to create one solution to make FAIR feasible.

Russom (2017)	W	Proposed a checklist for data catalogues in the digital enterprise. Emphasises that metadata is crucial and that employees must become data citizens.

Appendix C

Exhaustive literature research on associated topics

C.1 Data catalogues

Labadie, Legner, Eurich, & Fadler (2020) recognize opportunities between the use of the FAIR principles in an enterprise context and the use of a data catalogue. As explained previously, the data catalogue has the goal of maintaining an inventory of digital assets. This seems to be a good breeding ground for the introduction of the FAIR principles, they believe. As companies became more digital over time, the necessity to plan a data structure and accompanying architecture rose. As more data flooded the system, slowly it became possible to integrate data assets and associate them with business requirements. Enterprises started building data warehouses so that their data could be stored centrally, which enabled them to analyse their data more efficiently and expedite decision-making. As the amount of data grew, so did the need for accurate metadata, as companies realised their warehouses started to become cluttered. These were accompanied by business data dictionaries, that evolved from technical data dictionaries. The combination of the aforementioned components will then lead to a data ecosystem. This ecosystem consists of several, heterogeneous systems that store and analyse data. However, it is important that the data itself is treated and managed as if it were still in one system. This system can be referred to as a data catalogue, and it can be deployed to extract business value. A data catalogue has a set of 3 desired capabilities, but these might vary based on the user group.

- The data catalogue must facilitate inventory management of digital assets
- The data catalogue must have built-in features for collaboration for effective governance
- The data catalogue must have built-it features for communication about semantics

Important is that enterprises themselves take time to consider the strategy for access control since data protection is an ever-increasing concern.

As the implementation of FAIR will lead to an adapted way of working for many employees, there is a need for a change in mindset. This cultural and organisational

change must be managed effectively before employees will start to contribute actively to FAIR data repositories. To summarise, a data mindset must be developed, as well as a clear method on how the two concepts of FAIR and data catalogues can be combined.

Russom (2017) provides seven recommendations that are part of a checklist that will aid enterprises in selecting the location of a data catalogue. The seven recommendations will be summarized briefly to support the explanation of the data catalogue as done so above.

1. *Data cataloguing leads to an enterprise-wide view of data if relied on correctly.* This means that the catalogue must be seen as an inventory that is visible to data governance entities and stewards. The inventory allows for browsing through or searching for specific contents. Lastly, all the information should be in the catalogue, including metadata. This leads to the second recommendation.

2. *Metadata management must be modernised.* This means that the metadata must be made complete on three levels: technical, business, and operational. Metadata must be rich and added where possible, all while being connected to applications such as SAP or Microsoft. Lastly, the metadata, too, should be indexed in a catalogue that has various ways of entry.

3. *The data catalogue must contain several data types.* It must allow for the saving of unstructured data types such as Excel and Word files, but also machine-readable file types like JSON and XML. It is crucial that the catalogue allows for semantic search, as this will make the search process easier for the users.

4. *Tools should be included to handle the increasing amount of data, sources, and structures.* This means that data and metadata must be manageable through a variety of data management tools, where it is possible to recommend certain tools and interactions to users. The suggestion is that AI tools are used for the classification of sensitive data, anomalies in data, and correlations between datasets.

5. *Some users require other access rules than others.* Some functions in an enterprise will need the authorisation to alter datasets. It is important that this is possible without constant access requests that need to be checked.

6. *The data catalogue must allow for collaboration between different departments.* As many enterprises perform projects where the teams consist of employees from different departments, it is important that the data in the catalogue is accessible to all. The catalogue should allow easy data stewardship.

7. *The data catalogue should allow for automation of processes of data governance.* Data governance is an important aspect of a data catalogue. The catalogue must be compliant with regulations and must be secure so that private data is stored safely.

To summarize, data catalogues are a possibility for enterprises to store their data in an ordered, consistent manner. This is similar to the goals of the FAIR

principles. However, it can be a challenge to assign a clear checklist of steps to perform to ensure that data is produced and stored in a FAIR way (S. R. Wilkinson et al., 2022). The next section of the literature review focuses on tools that should make the management of FAIR less burdensome.

C.2 Change Management Frameworks

C.2.1 McKinsey 7S model

The McKinsey 7S model was developed by the McKinsey Company in the late 1970s. It is a conceptual framework that aids organisations in diagnosing the causes of their organisational problems and leading them towards improvement. The framework exceeds the classical notation that structure follows strategy, as it also incorporates the two with five other elements. The framework splits the seven elements into two categories: hard elements and soft elements. It is mentioned that hard elements are easier to manage than soft elements, as the latter are often less tangible and more influenced by company culture and individual behaviour.

The hard elements encompass the following elements: strategy, systems, and structure. *Strategy* is the plan of action when anticipating change. These actions must facilitate the enterprise to improve, or at least defend its competitive advantage. *Systems* refers to the procedures that are in place to ensure that the enterprise functions. *Structure* is the way the change is orchestrated. The organisation should focus on parameters that are crucial to the evolution of the enterprise. An example of this could be the degree to which decision-making is decentralised.

The soft elements are shared values, style, skills, and staff. *Shared Values* lay at the core of the 7S model. This element is sometimes also referred to as 'Superordinate goals". Coordinated goals and values are the foundation on which an organisation is built. These values define the enterprise's key beliefs, its business goals, and the essence of its corporate culture. In order to implement change successfully, organisations might need to change their values so that they can succeed. A large reason for the failure of implementation is the lack of focus on new shared values and having employees embrace these values. *Style* refers to the style that management utilises with regard to its employees. This involves the activities of the CEO and managers, and how they manage their people. The style of management gives away certain corporate values, such as market orientation or risk-taking behaviour. *Staff* shows the divide between hard and soft challenges. This is where morale and motivation, as well as the pay scale and bonus systems, come into play. The framework considers the staff as a pool of resources that need to be deployed and given attention to flourish. The role of management here is to hire and attract high-potential individuals who are given a clear career path. Lastly, *Skills* depend on the chosen strategy. This element evaluates what skills are needed to implement the change and make it a success (Waterman et al., 1980; Waterman, 1982; Channon & Caldart, 2015).

Below, a figure of the 7S model as proposed by Waterman, Peters, and Phillips, 1980, p. 18 (Waterman et al., 1980) can be seen as an example.

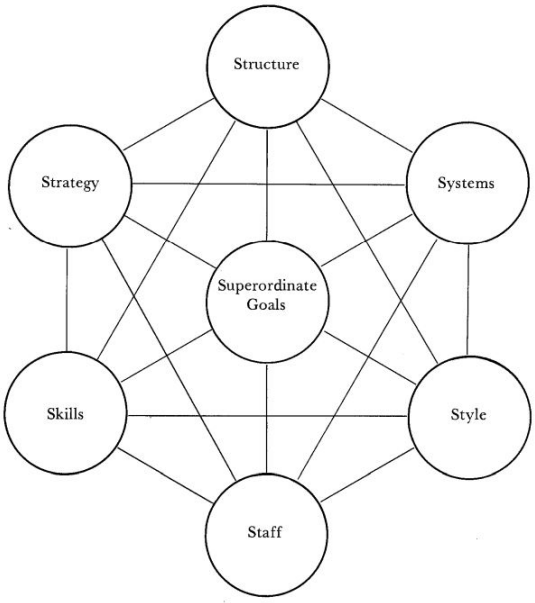

Figure C.1: McKinseys 7S model

C.2.2 Kotter

In 1995, Kotter published an article about why implementations fail. Here, eight errors were identified that caused the biggest problems (J. P. Kotter, 1995). Later, he published a book about leading change where eight steps were identified to support organisations in their itinerary to take control of their change management(J. P. Kotter, 1996). In 2014, a new version of the 8 steps of leading change was published. It is argued that the enhanced model better fits the velocity at which organisational change is required in today's times. Additionally, four Change Principles were designed to guide the process. In the figure below, the difference can be seen between the 8 steps of leading change and the 8 steps of accelerating change (Kotter, 2012).

The model for accelerating change is what will be used from now on, as the organisation at hand is subject to quick changes. The first step would be to create a sense of urgency. This is where project ambassadors are identified. These individuals must explain and describe the opportunity in such a way that it appeals to most individuals. Then, once employees are enthusiastic, the next step begins. A guiding coalition must be built using members from several layers of the hierarchy, preferably with different functions and networks. Together, this coalition conquers step three, where a strategic vision and initiatives are formed. These initiatives are activities that make the vision, a reality. Moving to step four, the coalition will recruit a large number of employees to help execute the strategic vision and initiatives. People will feel included and therefore are less likely to be resistant to the change at hand. With a significant amount of the enterprise on board, the barriers to action can be removed; step five. This can be in the form of eliminating inefficient processes or norms. These barriers can be identified through focus groups or workshops. They can often be found in the foundation of the enterprise, such as legacy rules and procedures, or the way the enterprise is organised on a technical level using silos. Then, after removing the barriers, there is room for small wins. These are crucial to the process as they will motivate people. These wins are to be communicated often and early to uplift the process of organisational change. Once a few wins have been

achieved, it is time for step seven. Here, acceleration is sustained. The project has increased credibility and can be used as leverage to create bigger successes. Change must be initiated until the final vision is achieved. This is where acceleration is key, as it is significantly harder to motivate a group of people after failing once. Then, the final step is step eight: institute change. The new vision must stick and become a part of the shared values in the corporate culture. The desired behaviours must be communicated, defined, and linked to business goals and success stories.

Figure C.2: Kotter's 8-step model for accelerating change (Kotter (2014) p. 9)

The four principles define four combinations that are key to transforming a business. First, Leadership + management. This is about motivating employees, executing a vision, and celebrating wins. Additionally, it concerns managerial processes like project management and reporting. The goal is to combine predictability and stability coming from management with the speed of innovation that is inspired by leadership to create a fast-moving, reliable organisation. Next, individuals require inspiration before putting effort into an idea. Just the data and logic will not be enough. There must be desire in order to take the organisation to the next level. This is what is incorporated into the principle of head + heart. Third, have to + want to. People must feel like they are involved before they want to participate in change. They know where the pitfalls are and they likely also know what the best way is to fix them. Lastly, some tasks in a transformation can only be carried out by a group of experts, however, it is impossible to leave all social aspects of organisational change to a select few. All levels of the organisation must be involved. The last principle refers to this under the title of select few + diverse many (Kotter, 2012).

C.2.3 ADKAR

The ADKAR model proposes two zones, enablement and engagement, and provides five key outcomes to ensure that a change is managed successfully. The first outcome, awareness, relates to the change asked of employees. They have to be informed before they are willing to make changes. This is where opportunities are identified and changes are communicated. Secondly, desire focuses on the desire to change once there is awareness, as these are in no way the same. They will engage once they see the positive effect of the change for themselves. Thirdly, knowledge. Employees have to be informed about what the changes mean to them. Here, they are trained to work with the new system in a safe environment. Then, ability differentiates knowing how to do something in theory from knowing how to do it in practice. This is where people are trained in the real-life system and where the gap between application and education is overcome. Lastly, reinforcement relates to reviewing the solution and whether it is fulfilling its expectation. Employees should receive coaching and there should be review sessions. This is where the change is sustained (Hiatt, 2006).

Appendix D

Case study: Coding

In the table below, the coding used for the case study results is shown.

Table D.1: Case study results related to current use and FAIR opportunities

Code	Sub code	Context
Current system	Benefit	Enterprise-wide
		Standardised structure
		Trainings are available
		Data leaks are minimised through deletion
	Challenge	No oversight
		Local machine storage
		Unclear method for storing and deleting
		Procedure slows down storage process
		Data silos
		Process is manual
		Gaps in documentation
		Not efficient as only management can store data
		Extraction from client system is difficult
		Intention to store drops when a project is done
		Latest version of files is hard to find
		Data is lost in folder structure
	Wanted change	Checklist for storing data
		Set folder structure
		More metadata
		Better reuse of data
		Search engine
		Tagging
		AI use
FAIR	Benefit	Speed
		Blacklist sharing
		Better communication
		Data sharing across daughter firms
		Scalability
		Confidence and comfort in data use
		Easy to understand framework
		Easily GDPR compliant

	Focus on metadata
	Smart indexing possibilities
Challenge	Shift from competition to cooperation required
	Must be enterprise-wide
	Data = power; no intention to share
	Workarounds are possible
	Challenging change management project
	Mindset must change
	Confidentiality is hard to ensure
	Strict access and authorisation rules
	Third parties will not be FAIR
	IManage compliance
	Creates more overhead and work
FAIR code	F1
	F3
	F4
	A1
	A1.2
	R1
	R1.3
Opinion	Clear
	Logical
	Makes sense
	Open door
	All for sharing
	Surprised this is needed

Appendix E

Requirements & suggestions taken from case study

This appendix contains requirements taken from the case study. First, a list will be provided that shows the requirements. Afterwards, the suggestions for the change management plan are shown.

The following requirements were identified from the results of the case study ordered by the requirements coming from the current system, followed by the requirements for the new principles (although some requirements might be applicable to both), and then the other requirements that are in neither category.

- Requirements from the current system
  - Version control
  - Guidelines/Checklist
  - Automation of tasks
  - Earlier iManage access
  - Security
  - Confidentiality
  - Increased findability
  - Increased metadata focus
  - More centralised platform for storage
- Requirements for the FAIR principles
  - Focus on internal application
  - Maintained data
  - Up-to-date data
  - Quality assurance
  - Clarity that open does not mean accessible to all
- Alternative requirements
  - Proper change management
  - Long-term focus

Table E.1: Overview of the answers from participants with relation to change management

Change management category	Factor
Social	Explain the need
	Make it easy to follow
	Take people into account
	Do not forget your people
	Do workshops, pilots, or focus groups
	Communicate timely
	Allow people to fail
	Ensure management support
	Have a change ambassador
	Use a project champion
	Do it for the 98% that is willing
	Provide training
	Take your time
Governance	Apply clear governance
	Provide long-term focus
	Create a change ecosystem
Technical	Ensure your technology works
	Be as bug-free as possible when starting

Appendix F

Relevant quotes from case study participants

F.1 Current system

The first question of the interview was related to the tasks the participants encounter in their job. These tasks are extremely varied, and the participants all come into contact with the data storage process in different ways. This led to all of the participants encountering different challenges in the way the process is currently orchestrated. Two participants noted that the storing of data on local machines is a challenging factor in ensuring that data folders are complete.

> *"[...] are stored somewhere and that somewhere is a shared folder structure and there's no real oversight over the whole folder structure. I think everything that is official, like KCW and iManage is all fine. But it's about like the whole Windows Explorer parts where people just save stuff and your e-mail boxes of course."*

The second participant added the following about the use of local storage:

> *"[...] also use quite a lot of other systems project systems to work from so Microsoft Teams is really a big one there. I also used myself because of this functionality and the integration with all the office tools. But let's say our policy is to only work from the iManage applications."*

This indicates that although a system is in place, not everyone uses it as the system is not that well integrated with other systems. A third participant adds the following:

> *"We store this on iManage. We store this on all kinds of team drives so normally. So I would say teams and iManage. And a little bit on local machines as well. [...] we assess what data we still need in teams, folders and all our local machines and what data can we actually delete because all that information on these contracts was all great to acquire we need that. But now the migration is done and it can actually be deleted. And wonder how strictly. That guidance is followed. [...] The typical pitfall there is that people store stuff on their local machines. And then you end up with "yeah, no, that's not the latest version I have that on my machine", yeah. And that's what you want to prevent. And I think that's why Teams is so great is that you can work in parallel in one file and it has version control and you don't have to mess around with local files. "*

Some managers indicate that there are strict rules and guidelines for storing data in the system, and find them easy to follow.

> *"There are a lot of trainings within ******* [1] internal trainings, so there is also one on this topic I think more on how to manage the engagement folder. So also how to deal with data would be one topic, but then eventually it comes down to the focus of the manager and senior manager in terms of the follow-up and you know how it is executed. "*

> *"[...] our policy is to only work from the iManage applications."*

> *"[...] but that's like a guideline, a policy. So it's not let's say that there are automated controls around it in our policy we say OK, you can only start [...]."*

Whereas others know that there are guidelines, but find it harder to follow them and would prefer a more framework-like checklist.

> *"[...] Maybe I really like a very clear list of what I need to do, and what is expected of me. Especially things that are not really about the engagement itself, but more about the tasks that need to need to be in order in order to do an engagement correctly. "*

One participant noted that their team created their own way of working with relation to both the guidelines and the storing on local devices.

> *"This is where you put everything, and that's something that you and the team decided, but that's not something necessarily every team does, right? [...] No, no, no. There was a particular in that team; our team. We decided we create one folder and then that's the place everybody start collecting it. But let's put it in that folder. So we all know where it is. And yeah, there was a team decision that project team decision. Exactly."*

Two participants noted that the current data process is not fluent as data saving starts before a project starts. Unfortunately, the codes used for the management of files do not get created before the project is officially started. One participant commented

> *"The only thing is that iManage works based on engagement code. But before you have an engagement code, there's already documentation that is or you already have documentation from clients, but also from other systems. [...], and you need to store that somewhere centrally. You cannot leave it in a mailbox. So you cannot use iManage. at that point yet. And I think that is also where people use and myself as well, like a sort of folder structure where you can keep all the relevant documentation before iManage or something else opens. [...] So it would really help If, if it's not engagement code specific, to be honest."*

Another participant added that many managers have found a workaround for this inefficient part of a system of engagement codes.

[1] ******* is used to blank out the company name

"So we have a yeah, strict guidelines. What should be filed. And that is that is quite strict and that's also used in in the different reviews and let's say audits, but that's like a guideline, a policy. So it's not let's say that there are automated controls around it in our policy we say OK, you can only start."

All participants agreed that it is beneficial that the entire enterprise uses the same system. One participant had an additional benefit to the way the process is currently designed:

"You limit any risks of data breaches. Think that is data leaks. That is what you want to prevent at any time and I would say that's the most important advantage that you don't retain any data that is actually sensitive but not needed anymore. [...] In the case that somebody enters the system, at least there are no privacy-sensitive files in there. [...] Or you can make a mistake yourself by sending a wrong file or whatever to a client that has client-sensitive information from another party."

This is important as security and confidentiality were mentioned several times across answer sets. When requesting participants to name something they would like to see added to the current process, the answers varied a lot. One participant would prefer a checklist, as mentioned before, whereas others are more concerned with the data itself and would like to see an increase in metadata use.

*"There is also a need for ******* to take their metadata more seriously [...] Yeah, that would really help. Not only let's say reuse it among the different communities. When I talk about tech and we don't, let's say event the wheel or store data twice and that we also let's say can easier identify data which should have the proper restrictions in terms of authorizations or confidentiality."*

Finally, one participant commented on the difficulty of finding documents back, and that a search engine with tags would be a good solution to this problem.

"A good search engine, because now it's always difficult. You need a code or you need something or a project name and a good search engine like for instance how Google works. I always like the Google search engine. You type something in a keyword and it finds it and I always see people complaining. If you have a, how would I call it a search filter, then you have 15 fields and you can search on these fields. Then it's mostly that 70% that people will complain about it, they will not find what I'm looking for. [...]. A system of keywords or perhaps tags for documents would be beneficial. "

After the questions about the current system, the participants were asked about the FAIR principles and whether they think that the principles are a solution to the challenges they experience in their tasks.

F.2 Answers on FAIR opportunity

First, participants were asked if they were familiar with the FAIR data principles. Most of their answers were similar, with one participant making the striking remark:

> *"Maybe you can give me a better explanation of what I found on Google, please. It is about accessible of how to access data and there's a nice explanation of the four principles. [...] I'm still figuring out what is meant by [interoperable]."*

, which was rather indicative of how the questions were answered thereafter. Most participants experienced difficulty with the definition of interoperable. This indicates that in order to make the FAIR principles usable for an enterprise, some definitions might have to be adapted. They all deemed the other principles to be clear and logical. One participant remarked that they were surprised that a framework for this process is still needed.

> *"My first reaction is: do we still have to think about this? Because how long do we use IT? already for many many, many years. Uh. And then we still think now about: how can we find data. Is it accessible? Can we reuse it and that, it's not that it worries me, but I I just think it's strange because we had so many systems, and tools. Ideas. search engines. Now we move to intelligence. Artificial intelligence. We have this chatGPT. It's crazy what is out there, outside. And then you think we still have to think about finding stuff. Is it accessible? Is it reusable? It's really strange, but apparently, it is needed."*

As can be seen in Appendix E, participants were able to identify a variety of benefits that enterprises could gain from implementing the FAIR principles. The first benefit that was mentioned, was the notion that it creates efficiency within a process, thereby speeding up that process.

> *"If you take universities as a research organization, I think they're very useful. You also saw it with the COVID medicine or vaccination. If everybody was together and information can be freely shared, it does actually have a lot of benefits in terms of speed."*

The same participant also noted that data is already being shared across firms, but that this currently is not an orchestrated process.

> *"I do think [sharing] already does exist, but it is more through the network that people have Themselves than it is through Information sharing by using system or whatever. [...] the only thing that Is orchestrated to say if ******* International, or Global, says "OK, we're Going to push this Information or the framework"."*

Finally, they also mentioned that the FAIR principles can be beneficial in scaling opportunities across firms.

> *"[...] does give a lot of scaling opportunities, of course, and I think also why would you try to invent the wheel again here if they did it in France already, for example."*

The next participant added more concrete benefits to the list, such as an increased focus on metadata, as quoted previously, and the fact that the FAIR principles are GDPR compliant and easy to understand.

"A framework which is also easy to understand, which with some principles you can always say keep in in the back of your pocket or in the back of your head to start thinking actively. How to operate and manage your data gets more and more important because of all the GDPR type of regulations and data privacy."

Lastly, they commented that the FAIR principles can help employees gain confidence and comfort in their data and knowledge of the data. This is important in an organisation's itinerary to become more data-aware and helps to establish a trusting bond with clients.

"[...]those principles also help you determine, let's say how to correctly and confidently and in comfort use your data."

An important aspect of the FAIR principles is the opportunity for tagging and indexing of data. This did not go unnoticed. Two participants quoted on the benefits of tagging and indexing.

"[...] an AI type of indexation and use of those, those language words will make like easier make our lives easier in the near future."

"And I think indeed using. clearly explicitly include an identifier of the data they described [...] maybe that's also a little bit indexing itself. I can imagine. You know, maybe adding labels to it in terms of what is the topic or what is in there. That's something that helps as well."

One participant noted that the use of FAIR among organisations would be a good opportunity to share data on companies that they have worked for. This could create the opportunity for a potential collaborative blacklist.

"Have a sort of a blacklist. Of you know, companies that build, I don't know, weapons that are, you know, in conflicted areas that we don't want to do business with. I can imagine something like that, [...] I can imagine that you want to share that information with each other. [...] corrupted regime for instance to prevent doing business with the wrong person."

A final benefit that was identified, was the notion that the FAIR principles could create an archive of opportunities and past projects so that these sources can be utilized when starting up new projects with the same client. After a few experiences, they acknowledge that currently, opportunity data is not always shared.

*"I think we have to share more between also the firms. Like, ******* we are a single firm, but also within the firm. [...] you reach out to a client and then you hear "Ohh, a colleague of yours was a couple of months earlier here" and then I heard this 3-4 times. So I think we have only these silos, these islands and everybody has a lot of information and data and that is so captured within that silo. It's a pity you can share it, You're more powerful together, but also it's very more professional."*

The next question was about challenges they expect in the possible implementation of the FAIR principles in their organisation. Many of the responses are closely

related to change management, which is why the choice was made to create a separate change management plan to accompany the FAIR framework. Two participants commented on the need for a shift in mindset. Currently, the information can be leveraged as a competitive advantage compared to your colleagues. The benefit mentioned in the previous paragraph about sharing data, therefore also becomes a challenge.

> *"I think sharing and making it accessible is very easy to do, but people should be willing also to share. Sometimes people see that having data, and information is also power, and if I share it I lose my power. So I think it's more on people if are they willing to share so that the firm is getting powerful and not only on the people's side. Especially when it is with another company [...] that's our competitors I think before you do that, I again, I'm for sharing. Yeah, we should more than I think, have more cooperative teams than competitors. [...] but then you are cooperating and together you're stronger"*

Taken from this quote before the FAIR principles can be implemented industry-wide, there must first be a shift in mindset in the sector. Enterprises must go from being competitors to cooperating before there is a willingness to share their valuable data. One participant shared worries about third parties and clients not being FAIR compliant and then receiving data from them. Additionally, they mention that it depends greatly on the type of contract whether data is shared and to what extent.

> *"[...] usually work together in let's say separate contracts. So when we work with another party, they're contracted by a client. We are contracted by a client and also the with the client discussed how data is shared amongst each other and that's then usually done through a let's say a client secure SharePoint or a kind of iManage type of application which resides with the client. In case when we work, let's say directly with another third party. Yeah, we have, let's say. Secure file transfers those kinds of things to share specific content, but. We. Yeah, in when we have parties working on their us so we hire them, then they get, they get access to let's say a specific environment where we can share data with them"*

One final remark made by one of the experts shows a part of the change management challenge. They remarked:

> *"Don't believe we need to implement everything. In our processes, it will also generate more overhead, right? It will, yeah. create more work really to maintain and implement. And also the disadvantage and the pitfall was always having nice ideas that eventually at some point people lose focus and it's not maintained anymore and then you could end up in an even worse situation."*

showing reluctance to change if it causes more work and if there is no motivation to keep up with the changes.

The third question was related to the FAIR principles they think are most important to their organisation. As mentioned previously, not one participant picked a principle that was categorised under interoperable. The principles that were chosen,

are F1, F3, F4, A1, A1.2, R1, and R1.3. Most of the participants could not give a clear explanation of why they thought those were the most relevant principles but noted that those principles are close to the current way of working. One participant gave an elaborate explanation of why they considered principles A1.2, F3, and R1.3 the most relevant.

> *"Let's say A1.2 which is about authentication or authorization. That's what's really key. The two also have to do the right authentication and access to the data in place. [...] And that's the R1, which I think is quite important and data switch described with the priority of accurate and relevant attributes. Because that also. Let's say we'll make our life easier. When we search for search and data or want to have a clear understanding of of the data. So it's let's say that's described in R1 but also in let's say when we talk about. The kind of labels. So let me look for that. So like F3. [...] R1.3: Each domain-relevant community standards. I think that's also quite important because we duplicate a lot of data. So now there it all has the same text and same namespaces associated, but it's the same kind of data and information, but then that that's a store it in a different type of files which lack the different attributes so it will be easier also to find relevant data reuse data when let's say there's relevant."*

Important to note is that next to interoperable, also principle A1.1 can be misinterpreted. One participant mentioned

> *"If I look at a A1.1. Protocol open, free and universally applicable and implementable suggests that data is sort of shared free across everyone and I don't think that is beneficial."*

It is therefore important in the framework to underline that open is not the same as free, and that access rules will still apply.

The final question in the FAIR phase of the interview was whether participants missed something in the current FAIR principles that they would consider relevant for an enterprise context. All participants related their answers to the quality and maintenance of the data. One participant related it to the risk of people changing the data wrongly, or not uploading the updated version back onto the database.

> *"My take is on on maybe adding another one that's more on let's say on the data processing and let's say changing the data. So when it's about how to access data, look at it. But how do you deal with them further so I'm able to download it, but then it's out of my secure environment and I can change it? [...] What are the protocols or principles when you want to or you can change that data? [...] But I think because of the importance of manipulating the data. , I will expect some more detailed principles around it. I can of course put it under A1.2, but I think it doesn't hold the importance of data manipulation."*

Another participant agreed with this statement and added that the assurance that the data is up-to-date is crucial.

"the maintenance part of it. But I need to find nicer words to make it like "FAIRy" or something. But more and that is what I mean. So it's more the maintenance of the data itself, right you can find It you can access it or not. You can reuse it or not. But it needs to be up to date as well. Also, confidentiality would be something that puts us at the front. I think from a Client perspective, it's actually pretty nice if you have an organization that also functions as one When you interact with it."

Another participant placed relevance on the correctness and quality of the data, which, according to them, can be achieved through updates and data cleaning.

"what is the correctness? How correct is the data? How old is the data I can see? "Oh, I see here some interesting, but maybe it's already old data. It's outdated". And then the truth is already different. So the correctness or the. How you say it, the accuracy of the data is also important. [...] Have also good quality so you have to clean it, and it needs to be accurate so that also the quality is of the highest standard"

F.3 Answers on change management

The final section of the interviews was related to change management. An overview of the factors influencing change management can be seen in table E.1. Here, a summary of the most relevant change management factors that were not mentioned in the literature will be given. For an overview of the synthesis between the literature and expert opinion on change management, see table H.1. The first suggestion was to incorporate employees into the change process. Although this is often mentioned in the literature, there is little notion of the use case that was presented by one of the participants, who suggested the use of workshops, focus groups, and pilots. The first two should be used to incorporate people into the process and make them part of the change, whereas the latter should be used to test the proposed change.

"There was extra training and we had pilots. So yeah, That also is part of the change management process for people themselves. How can you let people and their behaviour or reluctance to change? The adaptation can be difficult. [...] Once, a lot of bugs were still in the system. It wasn't really quite finished. The deadline got pushed. It costs us a lot of time from an engagement perspective and that did not benefit the process and acceptance at all"

Another participant added that a pilot provides the opportunity to start small and leverage small success through the use of project champions and change ambassadors.

"Piloting, just, you know, start small. Keep it simple. [...] making sure that on a small scale, it is a success and that you can leverage that."

The use of a change ambassador is mentioned more often. One participant commented on how to gain these ambassadors specifically:

"ask [focus group participants] if they can be ambassadors for you. They know other people. They were in your workshop, they hear things and

they also are your air to other people. So your information is bigger. [...] Change management, it starts when the consultant leaves because then people will have to work with this system, they're confronted with it and they have to do it and in the beginning, it's tough. [...] Let them be still half year be the ambassador and go every week to your user group and see how it's going and if there are complications or not or how to tackle [...] Call these ambassadors and ask how it is going."

Thereby also suggesting that the change ambassadors stay present once the project is finished. According to the participant, the change ambassador is needed the most at the beginning, when acceleration still needs to take place, and all the way at the end, when the change needs to stick.

It is also mentioned that an ecosystem for change has to be created before the change even starts. This can be done through active governance. This includes having conversations about the worries and struggles, but also setting up a support system, such as a change ambassador, and allowing people time to figure it out and fail because that is the best way to learn.

"There are a lot of opinions certainly on risk and quality within our firm and everyone's entitled to that, so it also takes quite some governance of making sure that these types of principles can be implemented [...] Change management is all about organisational change management and people and adhering to a different way of working or maybe they need to be organised differently, they need to have a different type of skill set or even. A different type of mindset. [...] We also talk about the ecosystem. We are a blue company. We like to have a structured approach. All kinds of really, well, framework type of steps. We're very good at that, but especially things like culture and changing the mindset, that's something you cannot easily make tangible. And that's where especially communities really help. [...] You need to take people by the hand, step by step. Allowing also them to learn and to fail. Because that's quite an important element when you also want to change your mindset that you have also the possibility to really learn. That means that you also need to fail."

The creation of such a change ecosystem fits well with the desire of another participant, who requested a more long-term view on what the change will bring to the enterprise.

"the best strategy is not really about getting people on board or explaining because it's all it has a clear view and we understand that that's all fine, but it's more about showing that this is indeed a more long-term plan and Sticking to it. [...] Many colleagues would appreciate long-term focus."

Lastly, an important piece of advice was given by one participant.

"the best point that you need to realize is that there will always be someone that isn't satisfied and cannot be satisfied. So you have to write it for like the 98% of the people that are understanding and OK and sometimes a little bit angry but fine, and not for the people that always have something.

You don't have to have something for every exception. It needs to work for the 98%."

Appendix G

Data preparation flow

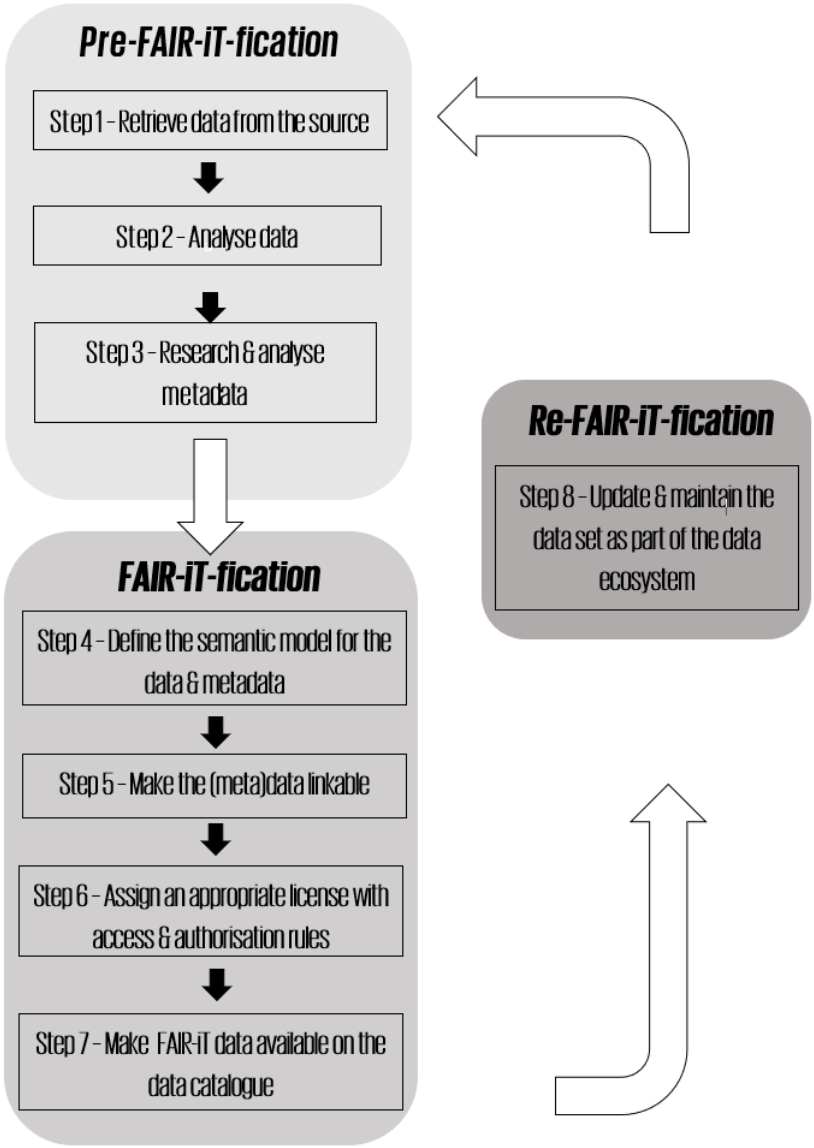

Figure G.1: The draft FAIR-iTfication process

Appendix H

Synthesis of literature and case study findings

Table H.1: Synthesis between literature and case study results

Concept	Literature	Author	Case study	Participant
Employee involvement	X	Akrong et al. (2022); Bharathi et al. (2012); Amoako-Gyampah & Salam (2004); Mesicek et al. (2021); Shah et al. (2011)	X	1, 3, 4
Effective communication	X	Akrong et al. (2022); Bharathi et al. (2012); Amoako-Gyampah & Salam (2004); Mesicek et al. (2021); Shah et al. (2011)	X	1, 4
Proper training	X	Supramaniam & Kuppusamy (2011); Tarhini et al. (2015); Alballaa & Al-Mudimigh (2011); Van Hau & Kuzic (2010)	X	2, 4
Explain the need	X	Tarhini et al. (2015); Shah et al. (2011)	X	1, 3, 4
Project champion	X	Supramaniam & Kuppusamy (2011); Alballaa & Al-Mudimigh (2011); Shah et al. (2011); Van Hau & Kuzic (2010)		
Top management support	X	Masheshwar & Javalagi (2019); Tarhini et al. (2015); Akrong et al. (2022)	X	4
Do workshops /pilots /focus groups			X	1, 3
Allow people to fail			X	2
Provide long term focus			X	3
Change ambassadors			X	2, 3, 4
Create a change ecosystem			X	1, 2
Minimize bugs			X	1
Appropriate governance			X	2
Plan accordingly	X	Tarhini et al. (2015); Shah et al. (2011)	X	2

Appendix I

Improved data preparation flow

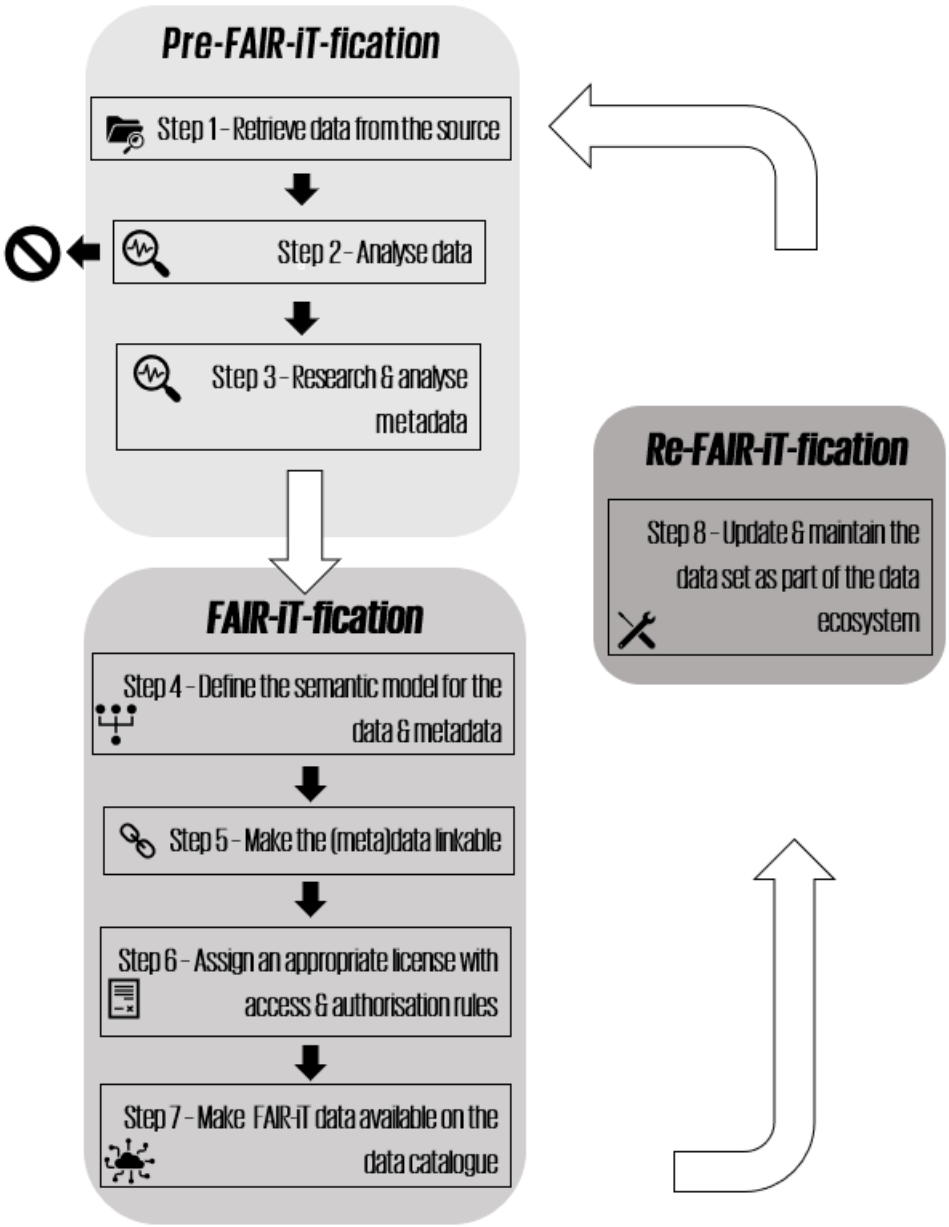

Figure I.1: The improved FAIR-iTfication process

Appendix J

The change management infographic

Figure J.1: The change management one pager

Appendix K

Validity and Reliability of Research

As proposed by Bryman (2012), research validation consists of validity and reliability. The *validity* refers to indications that indicators are designed to measure concept A really does measure concept A. *Reliability* refers to the consistency of the used measures. In the following sections, the case study results will be discussed according to Yin (2018). They propose four sub-types of reliability and validity for case study research: construct validity, external validity, internal validity, and reliability.

K.1 Construct validity

According to Yin (2018), construct validity refers to the establishment of correct operational measures for the studied concepts. This can be challenging in case study research as some subjectivity will always remain the in the data collection process. The initial measures collected during the case study in this research were focused on concepts such as requirements and the highlighted principles. These can be objectively measured and had the same meaning for all participants. Softer aspects are more difficult to measure and examples of this are change management effectiveness and challenges & benefits of several systems. These are difficult to quantify as they are dependent on the personal opinion and judgement of the interviewer and the participants of the case study.

It is suggested that construct validity is ensured by using various sources of evidence, establishing an evidence chain, and having participants review the published draft (Riege, 2003). The first aspect of collecting several sources of evidence can be considered achieved, albeit with certain limitations. All case study participants came from the same enterprise. This is not necessarily an obstacle as the framework is developed for that specific enterprise, but input from external sources would have been useful to make the framework applicable to a wider audience. A range of departments were taken into account when developing the framework, which should compensate for the evidence only coming from one enterprise. The second aspect was fulfilled as the evidence was transcripted and notes of observations are provided as evidence. The case study results were crosschecked with literary sources. The last criterion was fulfilled, but not fully, as some case study participants read their transcripts and one provided feedback on the developed framework. The preferred situation would have been that all participants could review their transcripts and

the artefact, but due to time, this was not possible. It can be claimed that overall, the criteria provided by Riege (2003) have been taken into account, although some aspects provide slight limitations to the construct validity.

K.2 External validity

External validity is related to the establishment of a domain to which the research findings can be generalised, according to Yin (2018). The research performed in the thesis has tried to create a framework that is applicable to a wider domain by viewing one of the principles as changing based on the developments in the sector. Multiple case studies using a replication logic should aid in increasing external validity, as has been done in the research at hand. Riege (2003) suggest comparing the collected data to evidence from literature, but this was not possible due to the lack of literature on this specific topic. The actions taken to increase external validity should make the findings generalizable beyond the performed case study, however, it is not fully clear whether this is the case. Yin (2018) mention that it is not possible to be fully certain of the external validity of research, even if several research methods have been applied in multiple contexts. The external validity of the case study at hand is limited by the number of enterprises in which the case study was performed, which is extremely small compared to the number of enterprises that could benefit from the implementation. Additionally, all case study participants worked at the Dutch firm, thereby creating a potential cultural bias. It is not possible to guarantee that external validity is fully fulfilled.

K.3 Internal validity

Internal validity is important to explanatory and causal case studies (Yin, 2018). The research at hand attempted to draw causal conclusions such as the implementation of X leading to benefit Y. Internal validity of these findings is of high importance. Yin (2018) suggests the following actions to increase internal validity: time-series analysis, explanation-building, and pattern matching. The first one has not been incorporated into this research. Explanation building is performed by providing extensive explanations of requirements and benefits and challenges. Pattern matching was done through a comparison of the case study results and identifying differences. Additionally, the collected data has been analysed using open coding techniques. This has provided both cross-case and within-case analyses of the collected data, which was suggested by Riege (2003). It can be concluded that the internal validity of the research has little doubt related to it, although it could have been improved through the use of time-series analysis.

K.4 Reliability

The reliability of research is characterised by the assumption that if the study would be repeated, the same results would occur. In order to ensure that this is the case, researchers should provide a case study protocol and create a case study database (Yin, 2018). The conducted research phases have been described as detailed as possible throughout the thesis. Due to the non-disclosure agreement and the right of

participants to remain anonymous, some factors will have to remain vague. It is not possible to describe the enterprise and departments the research was performed at in detail as this gives away too much information. This limits the replicability of the study. Even though all participants mentioned a different set of requirements, benefits, challenges, and experiences, a second execution of the interviews will likely lead to similar results as the research at hand, as most answers did have an underlying coherence. It is expected that the same change management steps will be mentioned as many participants have finished degrees that study this field, and a large portion of their answers stemmed from their knowledge of the topic. Furthermore, it is expected that the same challenges and benefits will be mentioned if the case study was performed again, as most participants had the same baseline of challenges they encountered.

K.5 Design research quality

The analysis that was performed in the previous sections was based on the reliability and validity criteria for case study research. As this is design research, the guidelines from Hevner, March, Park, & Ram (2004) will be used to demonstrate the validity and reliability of the design part of the thesis. They propose seven guidelines for design science research. When analysing the thesis, evidence can be found that all guidelines have been fulfilled. The most inept of the research is related to guideline 2. The solution is more management focused than technology focused. However, the presence of the data preparation model provides a limited amount of technical knowledge necessary to complete the steps. Additionally, guideline three could be improved through the addition of a naturalistic evaluation, instead of only expert opinion and case study respondents. This could be potential future research. Overall, the guidelines have all been satisfied to some extent and attempted with rigour, thereby suggesting valid and reliable design research.

Table K.1: Guidelines applied to the research according to Hevner et. al. (2004)

Guideline	Description	Application
Design as artefact	A viable artefact must be produced in the form of a construct, model, or method	The FAIR-iT framework is provided, including a data flow model and change management plan.
Problem relevance	The designed artefact is developed based on a technological solution with relevance to the business problem	Engagement data management has been identified as the relevant business problem. The solution strategy is both technology-based and organisational.
Design evaluation	The design artefact must demonstrate utility, efficacy, and quality via well-executed evaluation	The FAIR-iT framework was validated through interviews with experts and case study respondents. Empirical validation is yet to be performed.
Research contribution	The designed research must provide clear contributions in the areas of the artefact and design methodologies	The performed is one of the first to explore the topic of FAIR in an enterprise context. It belongs to a limited number of empirical studies on the domain of FAIR in a private context.
Research rigour	The design research must be done with rigour as it relies on rigorous methods for construction and evaluation	A systematic literature review was performed. Additionally, structured coding was applied to the empirical case study results.
Design as a search process	The search for an artefact requires using available options to satisfy the desired requirements within the laws of the problem environment	The design process was of an iterative nature to define the artefact with optimal utility for the scope
Communication of research	Design-science research must be presented in a way that is effective for management-oriented and technology-oriented audiences	The research has been presented to a group of IT audit and IT consultancy professionals as well as their managers. Ultimately, it will also be presented and defended at the university.