OCTOBER 2024

A ROADMAP TOWARDS INTEGRATED CSRD REPORTING

INCLUDING GUIDELINES AND
BEST PRACTICES

MSc THESIS – IRIS VAN HEIJNSBERGEN - BIT

UNIVERSITY OF TWENTE.
Flawless Workflow

COLOPHON

Master Thesis

Title:	A Roadmap Towards Integrated CSRD Reporting
Date:	October 2024
Classification:	Open

Author

Name:	Iris Marina van Heijnsbergen
E-mail:	i.m.vanheijnsbergen@student.utwente.nl
Programme:	MSc Business Information Technology
Specialization:	ICT Management and Enterprise Architecture
Faculty:	Electrical Engineering, Mathematics and Computer Science (EEMCS)
Institute:	University of Twente

Graduation Committee

Dr. Patricia Rogetzer

Email:	p.b.rogetzer@utwente.nl
Organization:	University of Twente
Faculty:	Behavioural, Management and Social Sciences (BMS)
Department:	Industrial Engineering & Business Information Systems (IEBIS)

Dr. Ir. Randy Klaassen

Email:	r.klaassen@utwente.nl
Organization:	University of Twente
Faculty:	Electrical Engineering, Mathematics and Computer Science (EEMCS)
Department:	Human Media Interaction (HMI)

Dr. Abhishta Abhishta

Email:	s.abhishta@utwente.nl
Organization:	University of Twente
Faculty:	Behavioural, Management and Social Sciences (BMS)
Department:	Industrial Engineering & Business Information Systems (IEBIS)

Geanny Amor dos Santos

Email:	geannysantos@flawlessworkflow.com
Organization:	Flawless Workflow
Position:	Chief Executive Officer (CEO)

Tom Behage

Email:	tombehage@flawlessworkflow.com
Organization:	Flawless Workflow
Position:	Managing Consultant

PREFACE

My initial reason for choosing the field of Business Information Technology is one that many others share: an interest in IT paired with a desire to maintain a human connection. Now, five years later, this thesis marks the end of my journey at the University of Twente. During my bachelor's, I shaped my interest towards the track of Enterprise Architecture and ICT Management, where my master's allowed me to explore certain topics in-depth and broaden my horizons.

I have always enjoyed helping people, whether as a teaching assistant or a consultant. So, when I discovered the new Corporate Sustainability Reporting Directive and the challenges it brings, it immediately sparked by interest. This graduation project has allowed me to combine academic knowledge with practice, and I am grateful for the opportunity to bring my ideas to life.

During the last five months I continued learning and developing my academic growth. The first to thank for this are my main university supervisors, Patricia and Randy. Working with them has always been a pleasure, whether as teaching assistant or through my own projects, and I am very grateful for the guidance, knowledge, and challenges they have given me. My thanks also go to Abhishta, whose outside perspective provided fresh insights on decisions and research directions I was deeply immersed in.

From the practical field, Tom and Geanny were with me on a daily basis as company supervisors. Their knowledge, feedback, and insights helped shape this project, with a special thanks to Tom for this careful review of my work. Additionally, several colleagues offered valuable insights, expert opinions, and suggestions that contributed to the final outcome.

I am also grateful to Anna, Bjorn, Corné, Inge, Kevin, Luc, Nienke, and Pieter for being open to share their practical experiences with the directive. Their concerns, challenges, and ideas were crucial input. I have enjoyed our conversations and learned a lot about the real-world impact of regulations.

Finally, I want to thank my friends and family for their unwavering support. My friends and housemates made my time at university unforgettable. I loved our study sessions, trips, parties, and adventures; all things that I will cherish forever. To my partner, thank you for your patience and understanding throughout all these years. And to my parents, I appreciate you reading my work and trying to grasp what I have been working on at university.

This thesis is the result of many dedicated hours of work and passion about the topic. I hope it contributes to novel insights and directions for the field. Above all, I hope it makes for an engaging read.

Iris van Heijnsbergen
De Bult, 25th of September, 2024

EXECUTIVE SUMMARY

The Corporate Sustainability Reporting Directive (CSRD) came into force in 2023, requiring large and publicly listed organizations to disclose details on their environmental, social, and governance practices. This amounts to over 1,000 data points such as emissions or employee retention, introducing a range of challenges for organizations as they work towards CSRD compliance.

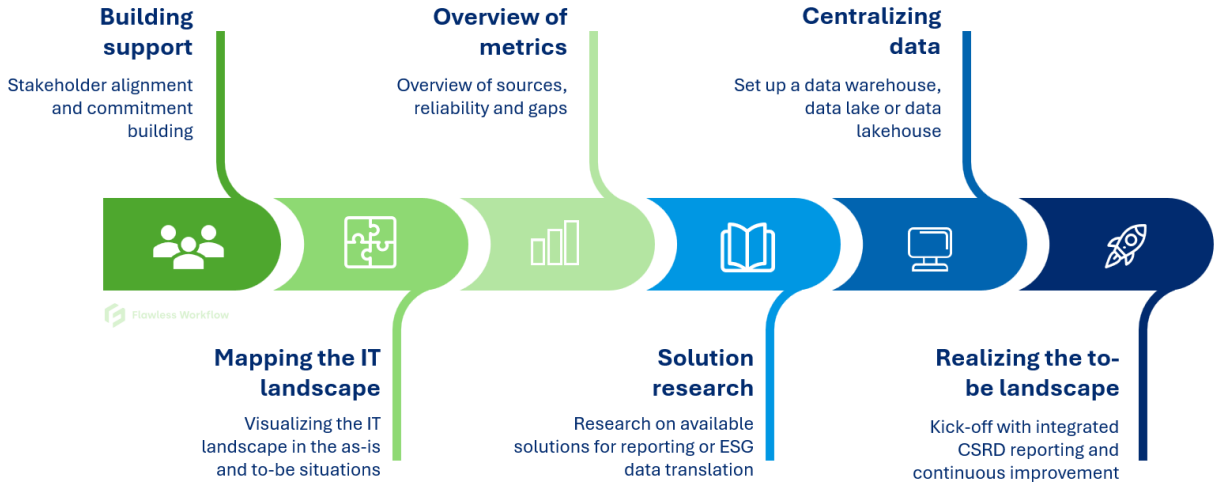
Among these, as identified in both literature and practice, data integration stands out as the primary challenge, particularly in terms of software connectivity, data collection and ensuring the quality and consistency of data. Although existing literature has identified these challenges, it falls short of providing specific solutions to address them. Similarly, while several practical roadmaps have been developed to guide general CSRD compliance, they do not offer the necessary steps for achieving integrated reporting. To address this gap, this research investigated the following main research question:

> *"How can companies integrate data from various sources within their IT architecture and along their value chain, facilitating data consolidation for CSRD reporting?"*

This question reflects a design problem, leading to the development of a roadmap and a set of guidelines aimed at supporting companies in achieving integrated CSRD reporting. These outputs help organizations gain control over their reporting processes, enabling a more streamlined and effective approach to CSRD compliance.

The research was guided by the Design Science Research Methodology as outlined by Wieringa, following the Design Cycle comprising problem investigation, treatment design, and treatment validation. A combination of systematic and exploratory literature reviews was conducted to define the CSRD landscape, identify relevant model-based analysis techniques, and explore data consolidation techniques. In addition, a series of interviews were conducted to gather empirical insights into the practical challenges and opportunities related to CSRD compliance.

Grounded in both literature and insights from the interviews, the created roadmap addresses key challenges and outlines best practices across six phases, offering a structured approach for effective data consolidation and integrated reporting. The high-level roadmap is presented in the figure below.



The first phase emphasizes building internal support and commitment, a critical factor identified by interviewees and supported by change management literature. Engaging with external stakeholders is equally important to ensure reliable data exchange across the value chain. The second phase involves mapping the IT landscape, drawing on enterprise architecture practices, to identify necessary technological adaptations. In the third phase, organizations are encouraged to document and evaluate the availability and reliability of data points needed for CSRD metrics.

The subsequent phases focus on solution research for reporting tools, data centralization via platforms like a data warehouse, lake, or lakehouse, and ultimately, the realization of an integrated IT landscape. By following this roadmap, companies can achieve effective CSRD compliance while enabling continuous improvement of their reporting systems. Therefore, this thesis provides valuable, practical insights that can be directly applied to facilitate the complex process of CSRD reporting within diverse organizational contexts.

The roadmap offers a valuable solution for organizations facing the complexities of integrated CSRD reporting. Validation through expert opinions and survey feedback confirmed its effectiveness, proving it can address key challenges and enhance reporting practices. The roadmap's adaptability makes it applicable across various sectors and organizational structures. Overall, the insights and recommendations presented in this thesis can serve as a critical resource for both future research and organizations that strive to streamline their CSRD reporting processes.

TABLE OF CONTENTS

1 INTRODUCTION 14

1.1 COMPANY INTRODUCTION 14

1.2 INTRODUCTION TO CSRD 14

1.3 RESEARCH DESIGN 15

1.3.1 Problem statement 15

1.3.2 Research objectives and scope 16

1.3.3 Research questions 17

1.4 STRUCTURE OF THESIS 20

2 RESEARCH METHOD 21

2.1 DESIGN SCIENCE CYCLE 21

2.1.1 Design problem 22

2.1.2 Research model 22

2.2 LITERATURE REVIEW METHODOLOGY 23

2.2.1 Exploratory literature review 23

2.2.2 Systematic literature review 23

2.3 INTERVIEW METHODOLOGY 26

2.4 INTERVIEW DATA ANALYSIS 27

3 PROBLEM INVESTIGATION – LITERATURE REVIEW 28

3.1 LITERATURE FINDINGS 28

3.1.1 CSRD 28

3.1.2 Model-based analysis techniques 38

3.1.3 Data consolidation techniques 41

3.2 DISCUSSION OF LITERATURE FINDINGS 46

3.2.1 CSRD reporting 46

3.2.2 Model-based analysis techniques 47

3.2.3 Data consolidation techniques 48

3.3 METHODS AND TECHNIQUES TO INTEGRATE VARIOUS DATA SOURCES 51

4 EMPIRICAL PROBLEM INVESTIGATION 52

4.1 INTERVIEW PROCEDURE 52

4.2 DATA ANALYSIS 53

4.3 INTERVIEW RESULTS 55

4.3.1 Drivers for motivation – Obligation, intrinsic, and opportunity 55
4.3.2 CSRD implementation – Timeframe, implementation process, vision, complexity, and transparency 57
4.3.3 Compliance factors – Double materiality and external challenges 60
4.3.4 Strategic considerations – Vision, central position, and business value 61
4.3.5 Stakeholder involvement and engagement – Stakeholders and internal challenges 64
4.3.6 Data infrastructure and systems – The landscape, source systems, software solutions, flexibility, and consolidation tools 67
4.3.7 Data points – Metrics and data challenges 70
4.4 DISCUSSION OF INTERVIEW RESULTS 73
4.4.1 Motivation 73
4.4.2 Implementation process 74
4.4.3 Technical considerations 76
4.5 CONCLUSION ON PRACTICAL CHALLENGES AND EXPERIENCES OF CSRD 78
5 ROADMAP DESIGN 80
5.1 PURPOSE AND SCOPE 80
5.2 DESIGN METHODOLOGY 81
5.3 ROADMAP PHASES 82
5.3.1 Phase 1 – Building support 82
5.3.2 Phase 2 – Mapping the IT landscape 83
5.3.3 Phase 3 – Overview of metrics 86
5.3.4 Phase 4 – Solution research 88
5.3.5 Phase 5 – Centralizing data 89
5.3.6 Phase 6 – Realizing the to-be landscape 91
5.4 PROPOSED ROADMAP 93
5.5 DISCUSSION OF ROADMAP 93
6 ROADMAP VALIDATION 96
6.1 EXPERT OPINION 96
6.1.1 Procedure 96
6.1.2 Expert opinion findings 96
6.2 SURVEY 100
6.2.1 Goals and structure 100

6.2.2 Conducting the survey 101
6.2.3 Survey outcomes 101
6.3 DISCUSSION OF ROADMAP VALIDATION 103
6.4 FINAL ITERATION OF THE ROADMAP AND GUIDELINES 105
7 GENERAL DISCUSSION 107
7.1 CRITICAL REFLECTION 107
7.2 LIMITATIONS OF THE STUDY 108
7.2.1 Interview sample 108
7.2.2 Roadmap requirements 109
7.2.3 Roadmap development 109
7.2.4 Validation 109
7.3 FUTURE RESEARCH 109
7.3.1 Validating the roadmap 109
7.3.2 Agile roadmap 110
7.3.3 Methodology for roadmap development 110
7.3.4 Application in alternative contexts 110
7.3.5 Digital tagging 111
7.3.6 OCR techniques and AI methods 111
7.4 RECOMMENDATIONS FOR FLAWLESS WORKFLOW 111
8 GENERAL CONCLUSION 113
8.1 REVISITING THE RESEARCH GAP 113
8.2 RESEARCH OUTCOMES 113
8.2.1 CSRD reporting 113
8.2.2 Applicable model-based analysis techniques in the context of CSRD 114
8.2.3 Data consolidation for CSRD reporting 115
8.2.4 Motivational drivers for software adaptation and digitalization 115
8.2.5 Implementation process for CSRD reporting 116
8.2.6 Technical constraints to achieve CSRD reporting compliance 117
8.2.7 Integrating data from various sources for CSRD reporting 117
8.3 CONTRIBUTIONS TO RESEARCH AND PRACTICE 119
9 APPENDIX A – SYSTEMATIC LITERATURE REVIEW 130
9.1 INCLUSION AND EXCLUSION CRITERIA 130
9.2 SEARCH TERMS 130

9.3 SEARCH QUERIES 131
9.4 DISTRIBUTION OF SELECTED LITERATURE OVER YEAR OF PUBLICATION 133
9.5 CONCEPT MATRICES 134
9.6 INTERVIEW QUESTIONS 136
9.7 CODING SCHEME 138
9.8 ESSENCE OF THEMES 140
9.9 DESIGN PROCESS 141
9.10 ROADMAP GUIDELINES 142
9.11 EXPERT OPINION QUESTIONS 143
9.12 SURVEY QUESTIONS 144

LIST OF FIGURES

Figure 1 - Problem cluster 15
Figure 2 - The Design Cycle by Wieringa (2014) 21
Figure 3 - Template for design problems by Wieringa (2014) 22
Figure 4 - Research model 22
Figure 5 - Numbers literature search and practical screen 25
Figure 6 – The phases of SSI guide development, by Kallio et al. (2016) 26
Figure 7 - Overview of the ESRS standards – adaptation of the researcher 30
Figure 8 - Overview of the GHG Protocol scopes by Bathia et al. (2011) 31
Figure 9 - Architecture Development Method by The Open Group (n.d.) 39
Figure 10 - Visualization of the data warehouse architecture 43
Figure 11 - Visualization of the data lake architecture 44
Figure 12 - Visualization of a data lakehouse architecture 46
Figure 13 - Visual comparison of the Data Warehouse, Data Lake, and Data Lakehouse by Lorica et al. (2020) 48
Figure 14 - Coding process 54
Figure 15 - Final thematic map, showing seven main themes 55
Figure 16 - Updated overview of challenges 79
Figure 17 - Roadmap for CSRD compliance by BDO (2024) 80
Figure 18 - Fictitious as-is EA model of an IT landscape 84
Figure 19 - Fictitious to-be EA model of an IT landscape 85
Figure 20 - Roadmap towards integrated CSRD reporting 93
Figure 21 - Revised roadmap guidelines 106
Figure 22 - Distribution of selected literature over year of publication, RQ 3 133
Figure 23 - Distribution of selected literature over year of publication, RQ 2 133
Figure 24 - Design Process by Munro (2016) 141

LIST OF TABLES

Table 1 - Key Performance Indicators for the Governance topic 33
Table 2 - Key Performance Indicators for the Social topic 34
Table 3 - Key Performance Indicators for the Environmental topic 35
Table 4 - Assessment of modelling languages 41
Table 5 - Comparison of the Data Warehouse, Data Lake, and Data Lakehouse 49
Table 6 - Demographics of organizations and interviewees; IP = In-person interview, O = Online interview via Teams 53
Table 7 - Suggestion for the overview of metrics, based on data points of ESRS E1 – Climate change 87
Table 8 - Inclusion and exclusion criteria 130
Table 9 - Search terms RQ 2 130
Table 10 - Search terms RQ 3 130
Table 11 - Search Queries RQ 2 131
Table 12 - Search Queries RQ 3 132
Table 13 - Concept Matrix RQ 2 134
Table 14 - Concept Matrix RQ 3 135
Table 15 - Codes per depth level, including their descriptions 138

LIST OF ABBREVIATIONS

ACID	Atomicity, Consistency, Isolation, and Durability
ADM	Architecture Development Method
AI	Artificial Intelligence
API	Application Programming Interface
BI	Business Intelligence
BMM	Business Motivation Model
BPMN	Business Process Modelling Notation
CSRD	Corporate Sustainability Reporting Directive
DBMS	Database Management System
DL	Data Lake
DMA	Double Materiality Assessment
DSRM	Design Science Research Methodology
DW	Data Warehouse
EA	Enterprise Architecture
EAM	Enterprise Architecture Management
EC	European Commission
ECB	European Central Bank
EFRAG	European Financial Reporting Advisory Group
ELT	Extract-Load-Transform
ERP	Enterprise Resource Planning
ESAP	European Single Access Point
ESEF	European Single Electronic Format
ESG	Environmental, Social and Governance
ESMA	European Securities and Markets Authority
ESRS	European Sustainability Reporting Standards
ETL	Extract-Transform-Load
EU	European Union

GHG	Greenhouse Gas
HR	Human Resources
IT	Information Technology
KPIs	Key Performance Indicators
LCA	Life Cycle Assessment
MDA	Model Driven Architecture
ML	Machine Learning
NFRD	Non-Financial Reporting Directive
NGO	Non-Governmental Organization
NPO	Non-Profit Organization
OCR	Optical Character Recognition
OLAP	Online Analytical Processing
PDF	Portable Document Format
RQ	Research Question
RTS	Regulatory Technical Standard
SA	Software Architecture
SBR	Standard Business Reporting
SLR	Systematic Literature Review
SMEs	Small and Medium-sized Enterprises
SoaML	Service-oriented architecture Modelling Language
SOE	State-Owned Enterprise
SSI	Semi-Structured Interview
SSoT	Single-Source-of-Truth
TOGAF	The Open Group Architecture Framework
UML	Unified Modelling Language
XBRL	eXtensible Business Reporting Language
XHTML	eXtensible HyperText Markup Language

1 INTRODUCTION

This chapter provides a comprehensive introduction to the foundation of this thesis. It starts with an overview of the company at which the research is performed in Section 1.1, detailing its core operations and positioning. Section 1.2 introduces into the context of the research: the Corporate Sustainability Reporting Directive (CSRD). The section highlights its significance and implications for the industry. Subsequently, the research design is discussed in Section 1.3, which includes the problem statement (Sub-section 1.3.1), research objectives and scope (Sub-section 1.3.2) and the research questions (Sub-section 1.3.3). Lastly, Section 1.4 outlines the overall structure of the thesis.

1.1 COMPANY INTRODUCTION

Flawless Workflow is a consultancy company positioning themselves as a strategic partner, specialized in digital transformations and creating an optimal workflow through the alignment of processes, software, and data. The company has a team that includes consultants, software developers, product designers, and workflow analysts. The focus of the company lies on technological improvements, while keeping the attention on the human side within organizations. With the strategic partnerships, there is a close collaboration between Flawless and their clients to achieve both short- and long-term goals with respect to digital transformation, ambitions and company growth [1]. Considering that clients and leads[1] start asking questions about CSRD implementation and data collection, Flawless identified the need for concrete guidelines on this topic.

1.2 INTRODUCTION TO CSRD

The Corporate Sustainability Reporting Directive (CSRD) was announced on the 20th of April 2021 and came into force on January 5th, 2023. The CSRD is designed in such a way that it supports the European Green Pact that aims to achieve net-zero greenhouse gas emissions by 2050, to become the first climate-neutral continent [2]. Furthermore, the strict reporting requirements should lead to a reduction in green washing, which is "*the corporate practice of claiming or exaggerating sustainability with the purpose of hiding a questionable environmental or socio-economic performance*" [3, p. 437]. The European Commission (EC) expects that companies must evolve and adapt their sustainability practices, which will ensure a more transparent approach towards emerging sustainability topics. This is further imposed through the requirement to disclose underlying reporting processes and governance structures [4].

The aim for the CSRD is to ultimately replace the Non-Financial Reporting Directive (NFRD) and solve the issues introduced with that directive [2], [5], [6]: the CSRD is said to be a revised and improved version of its predecessor [7]. The main weakness of the NFRD is that there is not one specific format required for disclosures, leading to companies choosing their own discretion while presenting relevant information about their sustainability

[1] Leads are companies interested in another companies' products or services.

endeavours [2]. Aboud et al., identify a similar issue to the absence of a format, explaining that there is a lack of clear guidance on how the directive should be enforced [8].

With a higher level of sustainability standardization introduced by the CSRD, several digitalization opportunities arise. As explained by Pizzi et al., new reporting tools can be developed, digital reports are generally more informative, and organizations can use digitalization to transition toward more sophisticated and reliable accountability approaches [9]. Digital technologies can make reporting tools more effective, with the potential to better track and report social and environmental conduct [10], as well as several data points related to governance. These opportunities are additionally noticed in practice, as 76% of organizations participating in the research by LeanIX indicated the importance of Key Performance Indicators (KPIs) and dashboards for tools supporting ESG initiatives [11].

1.3 RESEARCH DESIGN

This section presents an overview of the research design for this thesis. First, the problem statement is discussed in Sub-section 1.3.1, followed by a description of the research objectives and scope in Sub-section 1.3.2. The research questions are then presented in Sub-section 1.3.3.

1.3.1 Problem statement

Initially, consulting companies with a focus on accountancy (e.g., Ernst & Young, KPMG, and PwC) received questions from their clients about reports related to the CSRD. However, now more technology-oriented consultancy firms such as McCoy [6] and Flawless Workflow notice that their clients are increasingly asking about digitalization of sustainability reporting. Because of the changing regulations and growing number of companies that must comply due to new rules introduced by the CSRD, more and more insights arise into challenges they encounter. Organizations face challenges in understanding the requirements and determining the appropriate KPIs for the sustainability reports, while being unsure how to collect and integrate data coming from various data sources into one report [6]. Therefore, traditionally CSRD was an accounting issue, but nowadays it is increasingly seen as a digitalization issue.

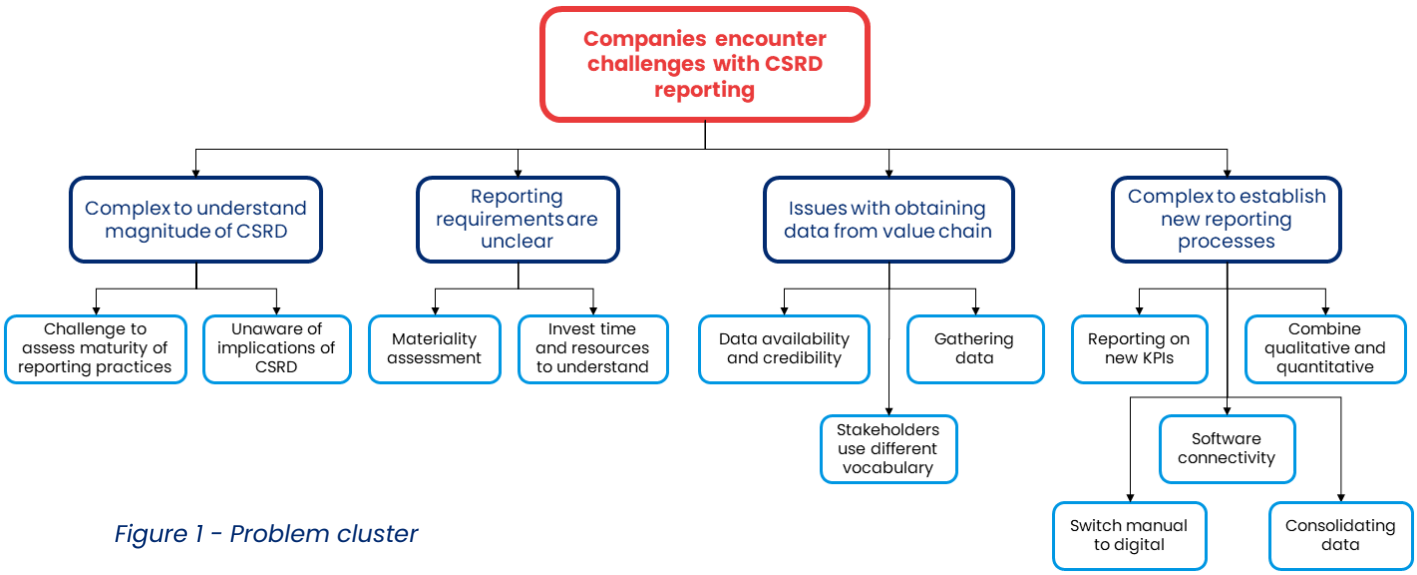


Figure 1 - Problem cluster

Furthermore, in literature challenges are identified related to changing traditional reporting ways [12]; making decisions regarding software and integration [13]; collecting data from the entire value chain [14]; combining qualitative and quantitative data [2]; and software connectivity [6] (Figure 1). These are all relevant considerations when having to combine data from various data sources. All in all, the main issue that has been identified is data integration, for which data collection serves as a sub-problem. These challenges span the categories of obtaining data from the value chain and establishing new reporting processes.

1.3.2 Research objectives and scope

Following the challenges outlined in Sub-section 1.3.1, a lack of literature in the CSRD domain has been identified. This is natural, considering it is a novel domain with the directive only being introduced in 2021. Due to the ongoing changes and developments in the field, much of the information does not accurately reflect the current state. Because of these disparities and an overall lack of knowledge, companies encounter several challenges with CSRD reporting. These are portrayed in the problem cluster of Figure 1 and are further explored in Sub-section 3.1.1. Much of the literature is focused on identifying these challenges, but there are no clear answers provided to solving them. Furthermore, several existing frameworks do consider the general CSRD process but fall short in offering specific steps to achieve integrated reporting.

Following these knowledge gaps, the objective of this research is therefore to provide a roadmap with concrete guidelines for companies on data integration and consolidation, such that they can gain control over their reporting processes. In doing so, several objectives are set out for this work:

- Provide an up to date and comprehensive overview of CSRD, with the intent that companies and fellow researchers have a solid foundation of the topic;
- Explore both model-based analysis techniques and data consolidation techniques, such that this knowledge base can be utilized for the roadmap;
- Conduct interviews with companies currently going through CSRD implementation, to identify best practices and specific challenges they encounter during this process;
- Design a roadmap with concrete guidelines that outlines an approach for data consolidation and integration for CSRD reporting;
- Evaluate the usability of the proposed roadmap and guidelines.

Additionally, the study provides an outline of companies' experiences so far, providing future research directions for the field. For the scope of this research, the focus for the empirical data collection lies on companies that are working to achieve compliance with the CSRD. This can either be through intrinsic motivation or because they are obliged and fall in the first or second mandatory group. This decision is made because the expectation is that these companies are fully embarked in the process of implementing CSRD, and employees therefore have knowledge of the challenges they encounter and best practices that have worked well so far.

1.3.3 Research questions

The overall research question (RQ-O) for this thesis is:

> *"How can companies integrate data from various sources within their IT architecture and along their value chain, facilitating data consolidation for CSRD reporting?"*

This question addresses a design problem, which results in the development of a roadmap and set of guidelines to support companies in achieving integrated CSRD reporting. To serve as a guide through this research, each chapter has its own main research question that serves as a sub-question to the overall research question. Chapter 3, the literature review, is the condensed version of a preliminary assignment of the researcher[2], for which the main RQ (RQ-LR) was formulated as follows:

> *"What methods and techniques can be utilized to integrate various data sources within companies' IT architecture and along their value chain, facilitating data consolidation for CSRD reporting?"*

This question aims to find both methods, such as methodologies and frameworks, and specific techniques, such as modelling languages and data consolidation techniques. To come to an all-encompassing answer to the main question, the research is divided into three additional research questions:

1. **What does CSRD reporting entail?**

The first research question is concerned with exploring the context of this research, providing details on CSRD reporting in general. As this is a relatively new field of research, the literature search is conducted in an exploratory form (Sub-section 2.2.1). Four pillars form the base to answer this rather broad inquiry. First, the key components of the CSRD are defined as an introduction to the topic. Next, as companies are struggling to understand the requirements, an overview is created based on literature and insights from the directive. Third, relevant KPIs are researched and listed; after which the focus lies on defining the main challenges that companies currently encounter with CSRD reporting. This will help identify the specific areas where this research can provide valuable insights and contributions.

2. **What model-based analysis techniques can be applied in the context of CSRD?**

Moving on beyond CSRD, the second research question focuses on the creation of a knowledge base concerning model-based analysis techniques. Numerous of these techniques can be used for different purposes related to IT architecture, such as creating a visualization of all data sources and how they are or could be connected. Understanding the IT architecture within the CSRD context is essential because it enables companies to comprehend their system's structure and the interconnections among data sources.

[2] To obtain access to the complete version of the Research Topics, please contact the researcher.

Therefore, it is valuable to identify how model-based analysis techniques can be applied to the CSRD reporting context.

For this question, a systematic literature review (Sub-section 2.2.2) is performed to create an outline of methodologies and approaches that characterize model-based analysis techniques. Modelling languages found in the literature review are evaluated based on four criteria: IT perspective, where a positive outcome means that the language can be used to model from an IT perspective; granularity, which can either be high (very detailed), medium, or low (very high-level); relations, where the language should have capabilities to model relationships between data sources; and diversity, which investigates whether a language is diverse or focused on one purpose.

3. **What are best practices for consolidating data from multiple sources for reporting purposes?**

Lastly, to aid companies with their main challenges, the final research question of the theoretical part aims to provide an outline of best practices for data consolidation. Several techniques are investigated to define the best practices in a reporting context. To create an understanding of existing approaches for data consolidation, an exploratory search is performed. Based on these insights, explanation of the techniques together with benefits and disadvantages can be presented.

Additionally, more details and information are gathered on data warehouses and data lakes as data consolidation techniques. Because data warehouses and data lakes are popular and well-researched techniques, a systematic literature search is conducted for this question as well.

Then moving from theory to the practical perspective, the following research question has been defined for the empirical chapter (RQ-EC), which will be answered through interviews:

> *"What are the practical challenges and experiences that companies encounter while implementing CSRD reporting, particularly in terms of motivational drivers and technical boundaries?"*

This question focuses on the practical implementation of CSRD, aiming to investigate both challenges and experiences. The latter can also be in the form of positive developments. Additionally, insights are gathered into the motivational drivers and technical boundaries that organizations encounter when it comes to CSRD implementation and digitalization efforts in this context. For more detailed insights, this question is also divided into three additional research questions (RQ 4 – RQ6), all addressed with the interviews:

4. **What are the motivational drivers for companies regarding software adaptation and digitalization in the context of CSRD reporting?**

RQ 4 delves into the motivational aspects that influence companies' decisions regarding the adoption of software and digital techniques for CSRD reporting. To gain deeper insights, this question is subdivided into the following:

a. *To what extent are companies motivated to commit to CSRD reporting and what are the drivers behind this motivation?*

The first sub-question (4A) explores the extent of companies' willingness to engage in CSRD reporting, examining the depth of their commitment and identifying the factors that motivate or discourage them from fully embracing CSRD efforts.

b. *How do the choices of companies for strategies and tools reflect their motivation and commitment to digitalization?*

Sub-question 4B further investigates how companies' choices in strategies and tools for digitalization reflect their motivation and commitment to digitalization of CSRD reporting. This will reveal whether there is an alignment between their motivation and the decisions they take.

5. **What is the implementation process that companies are going through for CSRD reporting?**

For RQ 5, the focus lies on the implementation process for CSRD reporting. Aspects that are researched in this case are the overall steps that have been taken towards compliance, stakeholders involved in the process, specific steps for the data consolidation process and future milestones that companies would like to achieve. Additionally, we will look at changes in the IT infrastructure that were or will be necessary to support CSRD reporting. Altogether, this will help us derive insights into the challenges they encountered so far as well as opportunities and best practices.

6. **What are the technical constraints companies face in achieving CSRD reporting compliance?**

This question examines the various technical challenges that companies encounter when striving to comply with CSRD reporting requirements. It delves into the specific types of data needed, the sources of this data, and the extent of its availability. By investigating these aspects, the research aims to identify the technical barriers that hinder effective CSRD reporting and understand how companies address these challenges. The question is further divided into the following sub-questions:

a. *What metrics do companies report on in a CSRD context?*

The first sub-question investigates the data required for CSRD reporting. This includes both metrics and KPIs that companies report on.

b. *From which sources is the required data obtained?*

Question 6B explores the data sources from which the data is obtained, both internal to the company and external from the value chain. This also involves the extent of data sourced from external parties within the value chain.

c. *To what extent is the necessary data available for companies?*

The third sub-question focuses on the availability of the required data. Companies may encounter issues with retrieving data from their own systems or with data coming from the value chain.

d. *What are challenges organizations encounter concerning data?*

Finally, this sub-question aims to find additional challenges that organizations encounter now that they have started the process, which perhaps have not been covered in literature yet.

1.4 STRUCTURE OF THESIS

This research is organized into eight chapters. Chapter 1 introduces the company at which this thesis is performed as well as the background information on the topic. Furthermore, the chapter outlines the design of the research, including the problem statement, research objectives and scope, and research questions.

Chapter 2 then provides details on the research method, outlining the steps of the design science cycle, the methodologies used for the literature review and conducting the interviews, and the techniques used for interview data analysis. Chapter 3 presents the results of the literature review, investigating the following areas: CSRD, model-based analysis techniques, and data consolidation techniques. Subsequently, Chapter 4 is dedicated to the empirical side of the problem investigation, addressing insights from practice through interview findings.

Chapter 5 addresses the main research question through the design of the roadmap. First, the purpose, scope, and design methodology are discussed. Subsequently, each phase is delineated in detail. The chapter concludes with the presentation of the final roadmap. Chapter 6 is then dedicated to the validation of the roadmap, through both expert opinion and a survey. Based on these findings, a revised version of the roadmap is presented.

Chapter 7 presents a discussion of the conducted research, incorporating a critical reflection, limitations of the study, avenues for future research and recommendations for Flawless Workflow. Lastly, Chapter 8 finalizes the work with a general conclusion as well as an outline of contributions to research and practice.

2 RESEARCH METHOD

This research involves two main components: descriptive research and design research. The descriptive research phase includes a dual approach of literature reviews and interviews, focusing on answering key knowledge questions that will inform the development of an effective roadmap. The project follows the Design Science Research Methodology (DSRM) by Wieringa [15], which is tailored for conducting design research while also addressing supporting knowledge questions. This methodology is discussed in Section 2.1. To address part of the knowledge questions, a literature review is conducted in both an exploratory (Section 2.2.1) and a systematic form (Section 2.2.2). For the other questions, empirical research is performed through a methodology for semi-structured interviews (Section 2.3) after which the collected data is coded and analysed (Section 2.4).

2.1 DESIGN SCIENCE CYCLE

In the DSRM, the primary objective of a design project is to (re)design an artifact to better align with achieving a specific goal [15]. In this case, the artifact has the form of a roadmap. Wieringa's DSRM is particularly suitable for developing a roadmap with guidelines for CSRD implementation because it emphasizes the iterative process of problem analysis, solution design, and rigorous evaluation. This approach ensures that the roadmap is not only theoretically sound but also practically viable and adaptable to real-world contexts. Previous research has similarly employed the DSRM as a framework for developing roadmaps [16].

The Design Cycle outlines the stages of such a research project, including phases such as problem investigation, treatment design, and treatment validation. This cycle is part of a larger Engineering Cycle that also encompasses treatment implementation and implementation evaluation. A treatment is *"the interaction between the artifact and the problem context"* [15, p. 28]. Depending on the outcomes of the treatment validation phase, iterations may be necessary until the designed artifact achieves the desired effects. The Design Cycle and the corresponding questions for each phase are illustrated in Figure 2 [15], where each bullet point indicated with a question mark is a knowledge question and the points indicated with an exclamation mark are design problems.

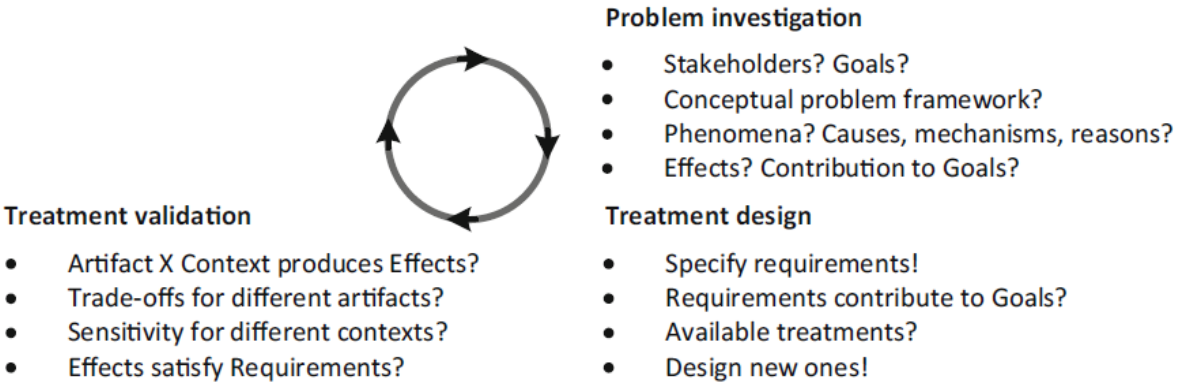


Figure 2 - The Design Cycle by Wieringa (2014)

2.1.1 Design problem

Wieringa defined a template for design problems, also known as "technical research questions" [15]. The template depicted in Figure 3 is used to formulate the objective for this research: improve *data consolidation for CSRD reporting* by designing *a roadmap with concrete guidelines* that shows *how to integrate various data sources* in order to help *companies gain control over their reporting processes*.

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

Figure 3 - Template for design problems by Wieringa (2014)

2.1.2 Research model

Combining the steps of the Design Science Cycle and the previously defined research questions, results in the research model as portrayed in Figure 4. This model is based on the notation developed by Verschuren & Doorewaard [17]. For each concept, the block in the upper left corner portrays the corresponding research question. The block in the upper right corner holds the reference to the respective chapter or section in which information is discussed regarding this question.

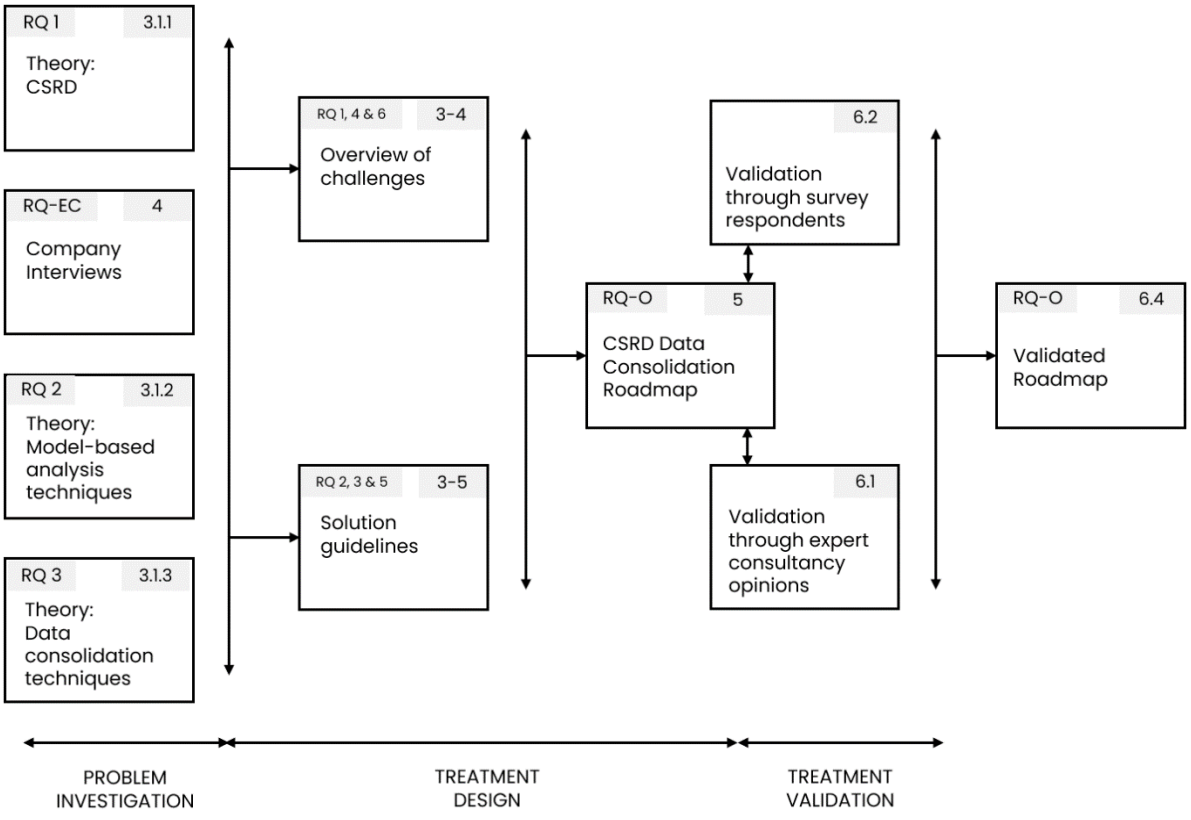


Figure 4 - Research model

2.2 LITERATURE REVIEW METHODOLOGY

The literature review is divided in an exploratory and a systematic part. This section is dedicated to outline the methodology used for each of them during the preliminary work of the researcher.

2.2.1 Exploratory literature review

The main part of the literature review consists of an exploratory literature search. According to Dash, such a search aims to explore the research questions while not intending to offer a final and conclusive solution to the identified problem [18]. The outcome is an overview of alternative options for a solution. For the research question on CSRD, we conduct exploratory search because of the novelty of the domain.

The exploratory search includes grey literature, for instance reports, working papers, government documents, master theses, and white papers. To the best of our knowledge, there are no predefined steps for this type of research, but the researcher has used the following guideline: first, identify the research questions; then, determine search terms based on key concepts in the RQs; after the database search, thematic analysis is performed; and lastly, the findings are synthesized. In doing so, there must be a focus on concepts instead of separate articles [18]. This guideline is, on a very high level, based on the steps of the systematic literature review as discussed in Sub-section 2.2.2.

2.2.2 Systematic literature review

Systematic Literature Review (SLR) is "*a systematic, explicit, and reproducible method for identifying, evaluating, and synthesizing the existing body of completed and recorded work produced by researchers, scholars, and practitioners*" [19]. Okoli & Schabram [19] developed an eight-step guide tailored to information systems research, drawing insights from Kitchenham [20] and Webster & Watson [21]. This guide is chosen for its suitability to the selected research questions, ensuring a systematic approach with the following steps:

Purpose

The first step is the clear identification of the purpose and intended goals of the review, in order to provide an explicit review [15]. Therefore, in the case of this research it can be defined that its purpose is to answer two specific research questions. These were focused on finding literature on the characteristics of model-based analysis techniques and on data lakes and data warehouses, including their use cases.

Protocol

The second step of the guide considers the research protocol and training of reviewers to ensure consistent quality, in case the work is performed in a collaboration [19]. However, as this review is conducted by one researcher this step will only consider the protocol. A protocol is "*a plan that describes the conduct of a proposed systematic literature review*" [20, p. vi].

First, the protocol considers the research question. In this case, RQ 2 and 3 are partly answered with a systematic review, the questions being:

- What model-based analysis techniques can be applied in the context of CSRD?
- What are best practices for consolidating data from multiple sources for reporting purposes?

For RQ 2, the systematic search aims to find the methods and characteristics that define model-based analysis techniques, creating a knowledge base that can be used to answer what specific technique can be applied to CSRD. The aim with RQ3 is more scoped as the SLR is focused on diving into the details of data warehouses and data lakes.

Then, the protocol defines the scope of the literature review, including search locations, screens, and keywords. In this case, the following digital libraries are used: IEEExplore, ACM Digital Library, ISI Web of Science, Google Scholar, ScienceDirect, and Scopus. To address the screens for each article, several inclusion and exclusion criteria have been defined for both research questions (Appendix 9.1). These will aid with effectively performing the systematic literature search [22] since they function as filters in the selection stage [19].

According to Wolfswinkel et al. [22], the set of search terms should reflect the entire scope of the research area. The search terms, queries and their variations per database are included in Appendix 9.2 and Appendix 9.3.

Literature search

After initial digital library searches for RQ 2 and RQ 3 using the defined search terms, "*it is important to supplement the search further to assure that all sources have been found and exhausted*" [19, p. 20]. This can be done using backward (snowballing) and forward searches, as recommended by [19] and [23]. Garousi et al. [24] suggest integrating published and practical work in fields like systems engineering. Grey literature such as reports, working papers, and government documents found through snowballing should be included, but quality assessment is paramount.

Practical screen and quality appraisal

After the initial search, articles are selected based on relevance and the criteria outlined in Appendix 9.1, which is done by reading the abstract of the articles [19]. Following Okoli & Schabram's advice to "*arrest first, ask questions later*" [19, p. 22], articles should be included if the reviewer is in doubt.

Quality appraisal includes checking the methodology, foundations for claims, and language. Wolfswinkel et al.'s method of 'refining the sample' [22] guides the process: removing duplicates, filtering based on title and abstract, and a full-text review. Additionally, for any interesting article forward and backward citations are checked. Based on these new articles, the cycle is performed again. This process yielded 18 articles for both RQ 2 and RQ 3 (see Figure 5): 12 from full-text review and 6 from backward and forward citation checks. The distribution of the selected articles over the publication years is portrayed in Appendix 9.4.

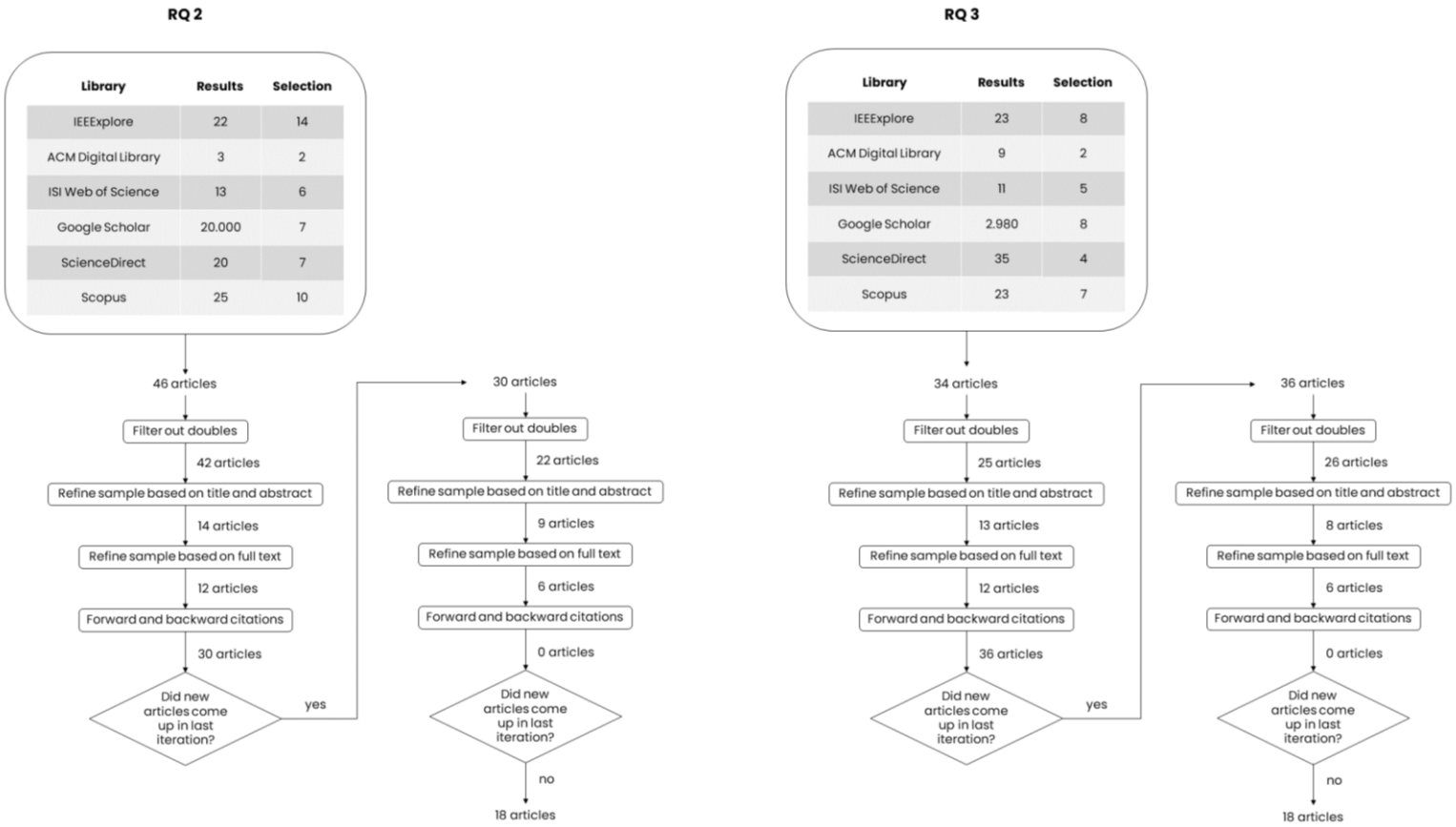


Figure 5 - Numbers literature search and practical screen

Data extraction

The data extraction phase, following Wolfswinkel et al.'s method [22], involves systematically gathering information from each article based on the defined research questions [19]. Starting with a randomly selected paper, relevant findings are highlighted and coded using open, axial, and selective coding methods based on grounded theory. Open coding creates categories from concepts and variables for an overview of the study. Axial coding refines sub-categories and identifies relationships, while selective coding integrates and refines main categories [22], [25]. This aligns with Braun & Clarke's thematic analysis phases [26]: familiarizing with data (in this case, literature), generating initial codes, identifying and reviewing themes, and producing the final report [19].

Synthesis of studies

The synthesis phase combines articles to create 'comprehensive sense' [19], transitioning from an author- to a concept-centric focus [21]. Using a concept matrix is recommended for proper literature synthesis, highlighting key concepts over individual papers [21], [22]. This aligns with Dash's observation that researchers may emphasize individual papers over synthesizing the question [18]. The concept matrices for RQ 2 and RQ 3 are included in Appendix 9.5. The final step then involves writing the review, reporting findings and relationships, and emphasizing novel insights; especially unexpected ones [19], found in Section 3.2.

2.3 INTERVIEW METHODOLOGY

A semi-structured interview (SSI) approach is selected for the interviews. Such an approach is employed when the researcher possesses sufficient knowledge about the subject to recognize its domain and primary elements but cannot foresee all potential responses [27]. The interviewees have the freedom to provide detailed answers, allowing for the exploration of a wide range of concerns regarding a problem [25].

In this research, the SSI is used in a descriptive/corrective manner, according to the typology by McIntosh & Morse [27]. For this interview type the aim is to confirm, refute, or elaborate upon the assumptions made by the researcher. The practical experiences of the participants act as a corrective to these assumptions [27]. SSIs are characterized by their design, where predetermined primary questions are followed by sub-questions or probes. The questions should be open-ended, such that discussion can be generated for which it is allowed to slightly diverge from the script. Probes can be scheduled as sub-questions but can also be improvised and arise from the dialogue. The main idea is that similar information is collected from each participant, by providing them with some guidance and gentle nudges whenever necessary [27], [28].

McIntosh & Morse [27] defined three steps for preparation of the interview schedule, including (1) identification of the domain of the topic; (2) identification of the categories under this topic; and (3) identification of the primary questions. The questionnaire should be critiqued and tested. Similar steps are outlined by Kallio et al. [28], as portrayed in Figure 6, and are used in this research.

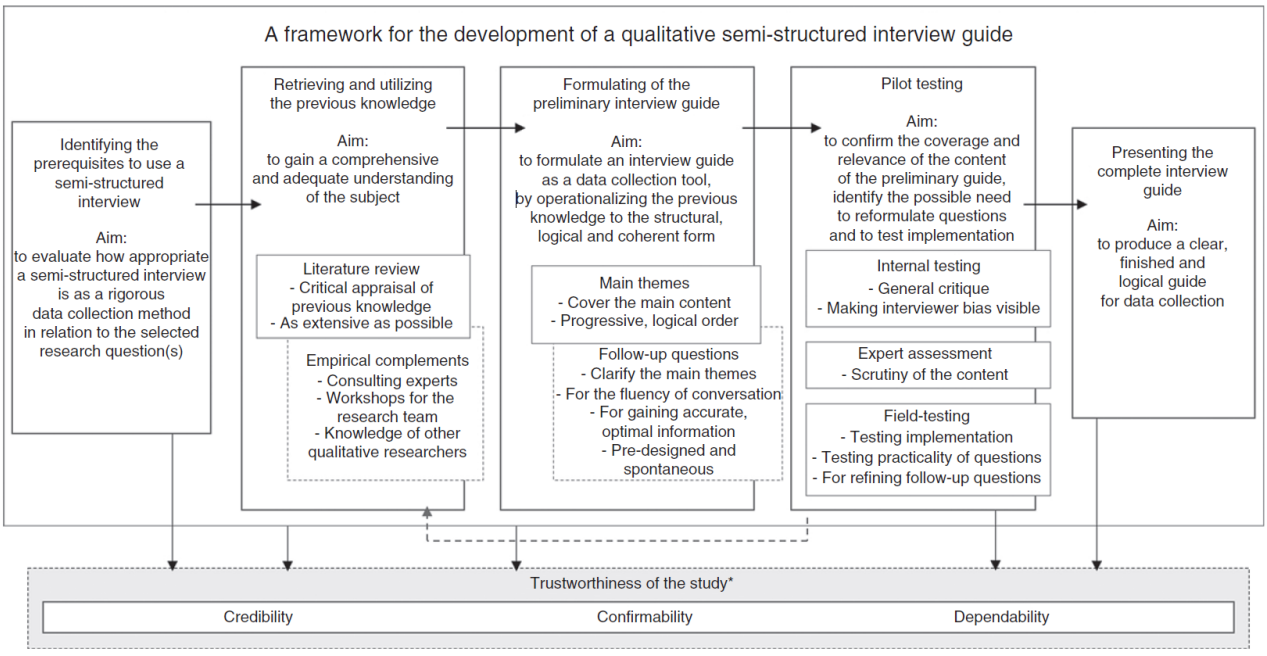


Figure 6 – The phases of SSI guide development, by Kallio et al. (2016)

Given that the primary language of the researcher and several interviewees is Dutch, most interviews for this research are conducted in Dutch. Following transcription, the transcripts are translated into English. Consequently, the quotes included in this research are direct translations from Dutch to English, except for those of interview 5 which was already fully conducted in English. The sample size of this research counts eight, and details about the interviewees are listed in Table 6 in Section 4.1.

2.4 INTERVIEW DATA ANALYSIS

The goal of SSI research is to understand participants' viewpoints to validate, adjust, or uncover novel insights related to the research topic. Consequently, the analysis of the collected data in this research offers a detailed and precise summary of participants' perspectives [27]. We follow the steps defined by McIntosh & Morse, which are data preparation and content analysis.

Data preparation includes transcription of audiotapes, which is done word-for-word without paraphrasing. To maintain confidentiality, certain identifying information concerning the participants is removed. Subsequently, content analysis then focuses on sorting and summarizing informational content of the data, by examining individual items and identifying common characteristics within the data. The first task includes deriving codes, which is done by item (question) [27].

The coding is performed according to the same steps as explained for the data extraction in the SLR (Sub-section 2.2.2). These include the phases by Braun & Clarke [26] and open, axial, and selective coding as described by Wolfswinkel et al. [22]. All in all, the content analysis comprises the stories as told by the interviewees into categories through emergent coding techniques [25].

Elaborating on the six phases [26], phase one includes familiarization with the data as well as transcription of verbal data. After reading through the data, phase two follows to generate initial codes. According to Braun & Clarke, "*codes identity a feature of the data […] that appears interesting to the analyst*" [26, p. 88]. These differ from the themes, which are often broader, that are sought for in phase three. In this next phase, the different codes are sorted and combined into potential themes. It is thereby important to consider their relationships. Subsequently, phase four includes reviewing and refining the themes. At the end of this stage, there is a clear overview of the different themes, their relation, and the overall story (Section 4.2 and Appendix 9.7). This is followed by defining and naming the themes in phase five, for which the scope and content of each theme is described in a few sentences (Appendix 9.8). Lastly, phase six involves production of the report, where the story of the data should be convincing to the reader. Here, extracts are embedded to provide narrative strength to arguments related to the research question (Section 4.3) [26].

3 PROBLEM INVESTIGATION – LITERATURE REVIEW

The following chapter synthesizes the findings learned through the explanatory and systematic literature search during the researcher's preliminary work (Section 3.1). These findings are the results of the thematic analysis, which was conducted according to a few steps. First, themes were identified while reading the articles. Second, for each paper, interesting points per theme were written down and included in the concept matrices. Lastly, based on these points the findings were synthesized into main categories, each focused on one of the research questions. Based on the findings, the research questions are answered in the discussion (Section 3.2). Then finally, Section 3.3 forms a conclusion and summarizes the findings by providing the answer to the main research question of this chapter:

> "*What methods and techniques can be utilized to integrate various data sources within companies' IT architecture and along their value chain, facilitating data consolidation for CSRD reporting?*"

3.1 LITERATURE FINDINGS

As mentioned above, each section discusses the main categories for each of the research questions. These are focused on CSRD (Sub-section 3.1.1), model-based analysis techniques (Sub-section 3.1.2), and data-consolidation techniques (Sub-section 3.1.3).

3.1.1 CSRD

Since the context of this research is tailored to the CSRD domain, it is important to obtain a proper understanding of this topic. Therefore, the first research question is determined as "*What does CSRD reporting entail?*". This sub-section focuses on defining the key components of CSRD; defining reporting requirements (including digital tagging, external audit, and sector-specific standards); investigating relevant KPIs based on the provided standards; and delivering an outline of challenges that companies currently encounter with the reporting, as well as their motivation towards CSRD in general and digitalization.

Introduction to CSRD

The CSRD entered into force on January 5th, 2023, requiring all large and listed companies to report on their sustainability endeavours. This new directive aims to modernise and strengthen rules concerning social and environmental issues, and companies' impact on people and the environment [29], [30]. In doing so, the hope is that companies can achieve sustainable and inclusive growth [31] while being transparent about their undertakings.

To ensure transparency, the sustainability reports will be freely available to the public, either on the websites of the companies themselves, through central registers, or through the registers of companies of the EU Member States [14]. Additionally, since the CSRD mandates that sustainability reporting be integrated into the management report, it finally places sustainability reporting on an equal footing with financial reporting [7]. So far, sustainability reports were separate and not mandatory at the same level of the CSRD.

Where the NFRD concerned public-interest entities, the CSRD will cover all large companies that meet at least two of the following three criteria [4], [6], [30], [31]:

- a net turnover exceeding €50 million per year;
- a balance sheet total of more than €25 million;
- more than 250 employees.

The estimation is that more than a thousand companies in the Netherlands will have to comply [30], with some having expectations that the number lies between 3,000 and 6,000 [32]. In the EU, the number will increase from 11,000 companies complying with the NFRD to 50,000 complying with the CSRD [5], [7], [14], [33].

The directive will be phased in for different categories of companies. Large, listed[3] institutions with over 500 employees, already under the NFRD, must start their reporting from the 2024 financial year. Large companies start from 2025, Small to Medium-sized Enterprises (SMEs) from 2026, and non-EU companies with more than 150 million turnover in the EU and a subsidiary or branch in the EU have to start from 2028 [14], [30], [31], [32]. As of now, the LSME and VSME[4] have been developed for public-interest SMEs and for voluntary reporting, respectively [34].

The European Financial Reporting Advisory Group (EFRAG) developed the European Sustainability Reporting Standards (ESRS) for CSRD reporting. The group is a collaboration between several national and European stakeholders, including Business Europe[5], the European Central Bank (ECB), Non-Governmental Organizations (NGOs) and trade unions [32].

The standards were published on the 22nd of December in 2023 and are tailored to EU policies [29]. They include two general standards: ESRS 1 – general requirements and ESRS 2 – general disclosure, mandatory for all companies. Additionally, there are eleven topical standards covering the ESG themes: environment, social and governance. Reporting on these themes is required if deemed material (see page 32); otherwise, companies must explain their materiality assessment conclusions [12], [31], [35].

Reporting content

Before diving into the standards provided by EFRAG, it is important to note that they focus on the responsibility of all companies within a value chain. Hence, companies must report on the ESG performance of their customers and suppliers (downstream and upstream) as well as on their own. Recognizing the complexity of retrieving information from the value chain, the EU allows an exception for the first three years if companies cannot obtain all value chain information. However, companies must explain their efforts, constraints,

[3] Listed companies on the stock exchange.
[4] ESRS for listed SMEs (LSME) and voluntary reporting standard for SMEs (VSME)
[5] Confederation of European Business

reasons for information gaps, and future solutions [31], [36]. EFRAG provides a list of steps for organizations to consider during this period, which include:

- engaging with stakeholders and enhancing the materiality assessment;
- preparing the necessary technological and other infrastructure for reporting;
- updating contracts with value chain actors to include new policy implementations or target tracking;
- enhancing understanding about the structure of the value chain, the actors involved, and the associated impacts and dependencies [36].

The value chain, as defined by EFRAG, is "*the full range of activities, resources and relationships related to the undertaking's business model and the external environment in which it operates*" [36, p. 8]. This includes all activities, resources, and relationships of an organization. The reason for including the entire chain is that the major impacts, risks, and opportunities of an organization mostly occur in its upstream or downstream value chain, rather than in its own undertakings; the same holds true for emissions. The aim is therefore to provide a complete picture of the impacts on people and the environment, along with proper identification of risks and opportunities.

As mentioned, EFRAG created two general standards and an additional set of ten topical standards covering the ESG topics [32]. Figure 7 provides an overview of these topics based on [30] and [6].

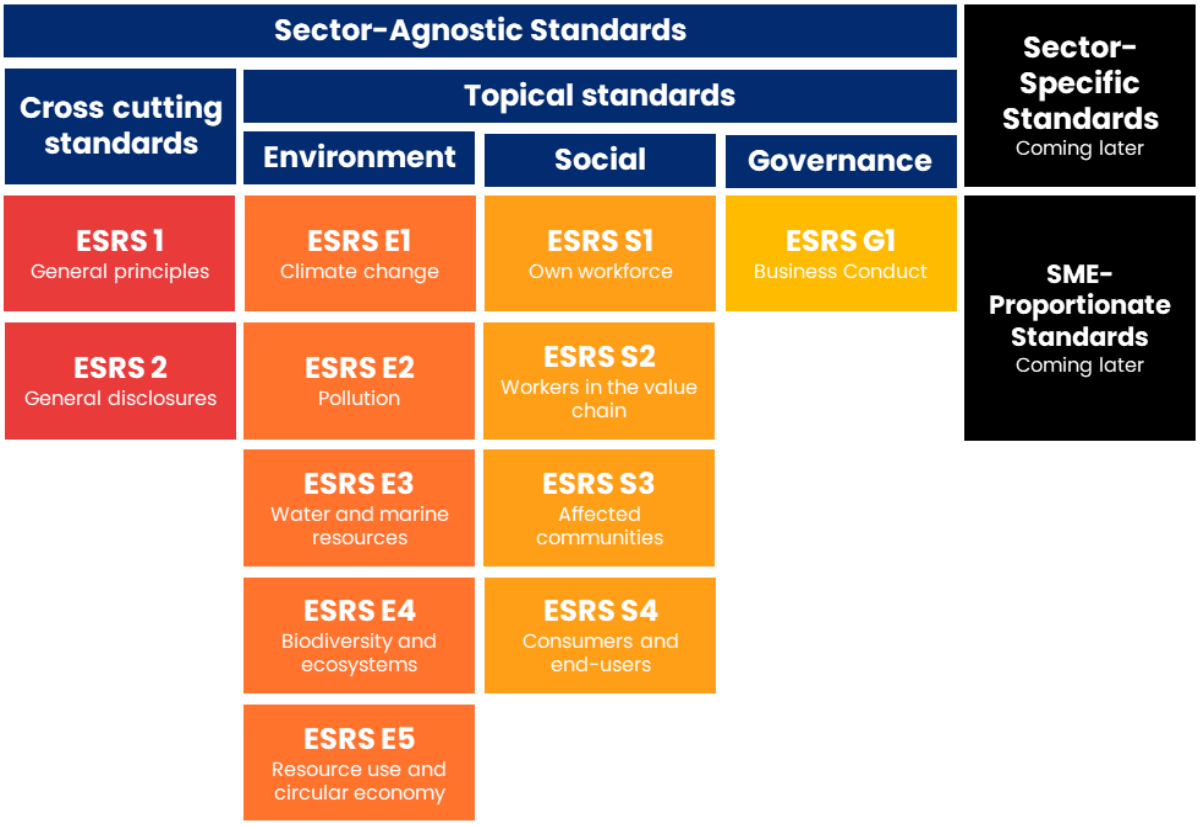


Figure 7 - Overview of the ESRS standards – adaptation of the researcher

In the context of the CSRD, measuring Greenhouse Gas (GHG) emissions along the value chain is critical. The GHG protocol divides this into three scopes. Scope 1 includes direct emissions controlled or owned by an organization; Scope 2 covers all indirect GHG emissions from purchased utilities like heat, electricity, steam, or cooling [37]; and Scope 3 encompasses all up-and downstream activities. Greenhouse Gas Protocol [38] created a clear visualization of the concept (Figure 8). Generally, measuring and reporting emissions along the entire value chain is a difficult aspect of reporting.

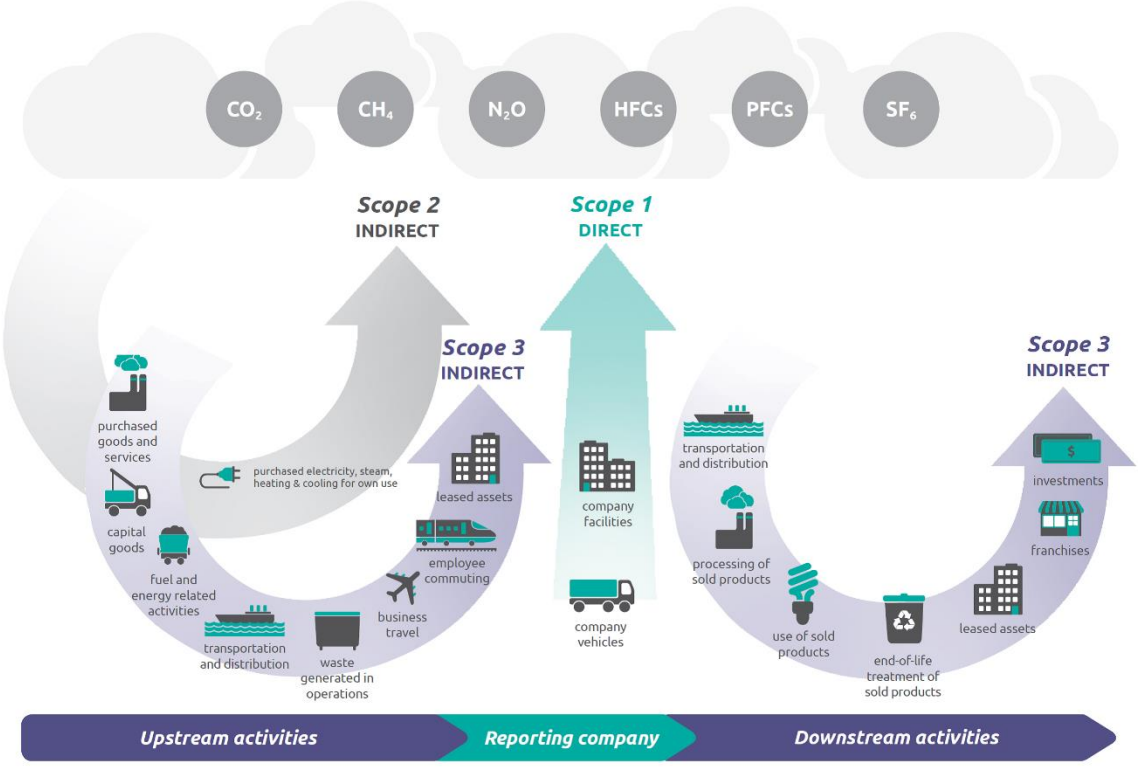


Figure 8 - Overview of the GHG Protocol scopes by Bathia et al. (2011)

The first set of ESRS, the sector-agnostic act, was adopted by the EC on the 31st of July in 2023 [39]. Sector-specific standards, initially expected in October 2023 [6], have been postponed (as depicted in Figure 7) to early 2025 for the oil and gas, mining, quarrying, and coal industries. For financial institutions, the standard-setting research process was started in the last quarter of 2023 [40] and workshops are held to collect feedback on the current version of the draft [41].

The actual adoption of the sector-specific ESRS has been delayed from mid-2024 to mid-2026 to reduce administrative burdens for companies and to cut the reporting requirements by 25%. The Commission understands the importance of a system that ensures reaching objectives at minimum costs, so organizations can now first focus on compliance with the sector-agnostic standards. The standards for non-EU companies have also been postponed to mid-2026 as their reporting obligations will only start to apply in 2028 [41], [42], [43].

The sector-specific ESRS should provide clear guidance on the level of detail and type of information that organizations in certain sectors need to disclose about their impact on people and the planet. As methods and impacts can vary by sector, this should include decarbonisation, biodiversity, and human rights. The standards are seen as a valuable source of information for investors, as they allow for comparisons between companies [43].

Materiality

The CSRD incorporates a two-sided perspective called double materiality, addressing both financial materiality (outside-in, how external ESG factors impact the company) and impact materiality (inside-out, how the company impacts external ESG factors). Financial materiality concerns the financial effects of sustainability risks and opportunities, while impact materiality focuses on the company's impact on people and planet [5], [12], [31], [32], [44], [45].

A mandatory materiality assessment helps companies determine relevant topical standards for their reports [6], [31], [35], [46], considering both their own operations and their upstream and downstream value chain [44], in the short, medium and long term [4]. Simply put, information is considered material if its absence or misrepresentation could sway the user's judgement. All material topics, whether from a financial or impact perspective, must be included in the sustainability report [32], [45].

Stakeholders' input is crucial for the double materiality assessment. The ESRS distinguishes between affected stakeholders (those impacted by the company's activities) and users of sustainability statements (primary users of financial reporting and other stakeholders) [39], [47]. Ultimately, the double materiality assessment is said to serve as a proxy for a much more fundamental debate on the social responsibility of business [46]. Organizations are required to disclose their materiality assessment process and its outcome, including methodologies, assumptions, focus, extent, and inputs [44].

Reporting requirements

One key CSRD reporting requirement is digital tagging. Annual accounts and reports must comply with the Standard Business Reporting (SBR), and CSRD adds another layer [31]. EFRAG is developing a digital taxonomy that reflects the ESRS in an eXtensible Business Reporting Language (XBRL) format, enabling organizations to tag their sustainable statements in a structured, machine-readable data format [9], [48]. This allows information to be analysed without manual and error-prone transformation from files in portable document format (PDF) or printed documents. According to EFRAG's implementation guidance "*the digital tagging needs to be performed in the European Single Electronic Format (ESEF) according to an adjusted Regulatory Technical Standard (RTS) that will be developed by ESMA, on the basis of the technical advice on the draft ESRS-XBRL Taxonomy that will be released by EFRAG in 2024*" [48, p. 5].

The process uses eXtensible HyperText Markup Language (XHTML)[6] with XBRL tags, enhancing accessibility, availability, and usefulness, and reducing administrative burden [2], [9]. Tagging, which labels data for automated reading and processing through machine algorithms [2], [32], optimizes data management and systems within organizations. This can for example be through software that assists in mandatory tagging through APIs or organizing data for CSRD publication [32]. Ultimately, the sustainability information can be effortlessly incorporated into the European Single Access Point (ESAP) [31], a centralized, freely accessible database [49].

The sustainability report must be integrated into companies' required annual management reporting and include an assurance statement issued by an external auditor. This should increase the quality, reliability, transparency and comparability of the sustainability information, bringing it to the same level as financial information [5], [30].

Compliance with the CSRD is mandatory; non-compliance identified by an external audit can lead to administrative sanctions and three potential penalties could be faced: public denouncement, an order to alter behaviour, and a financial penalty. Each EU member state will determine the penalty and set the boundaries for sanctions within their jurisdiction [6].

According to the Supplementing Directive of the ESRS, organizations do not need to disclose classified[7] or sensitive information, even if deemed material. However, they must still ensure the overall relevance of the disclosure, even when classified or sensitive information, intellectual property, know-how, or innovative results are omitted [35].

Key performance indicators

An overview of relevant KPIs for the CSRD reports is provided in Table 1(Governance), Table 2 (Social), and Table 3 (Environmental), but more are included as metrics in the Draft List of ESRS Data Points [50].

Table 1 - Key Performance Indicators for the Governance topic

Topic	Category	KPIs
Governance	G1 – Business conduct	Numbers of risks or complaints
		% Functions-at-risk covered by training programmes
		% Sustainable involvements
		Number of convictions or fines for violations
		Number of corruption or bribery-related incidents
		Financial political contributions
		Amount of internal and external lobbying expenses
		Average days to pay invoices
		% Payments aligned with standard terms
		Outstanding proceedings for late payment

[6] XHTML is a markup language designed for hypertext.

[7] EU classified information as defined in Council Decision 2013/488/EU on the security rules for protecting EU classified information or classified by one of the Member States and marked as per Appendix B of that Council decision [35].

Table 2 - Key Performance Indicators for the Social topic

Topic	Category	KPIs
Social	S1 – Own workforce	Gender diversity over current & past years
		Nationality within the company
		Training hours provided to employees
		Incidence of injury and sickness
		Non-employees in own workforce
		Employee turnover
		Numbers in age-ranges
		Disabilities
		Family-related leaves
		Renumeration ratio
		Complaints & human rights issues
	S2 – Workers in the value chain	Equal chances & discrimination
		Human rights issues
		% Permanent contracts
		Gender wage disparity
		Supplier evaluations (sustainability)
		% Employees living beneath living wage limit
		General living wage & working hours
		Forced labour & child labour
	S3 – Affected communities	Involvement in community initiatives
		Involvement in charitable actions
	S4 – Consumers and end-users	Impact and satisfaction levels
		Number of complaints received from consumers and/or end users

Table 3 - Key Performance Indicators for the Environmental topic

Topic	Category	KPIs
Environment	E1 – Climate change	GHG emissions (1, 2 & 3)
		Energy consumption
		% Renewable energy used
		Fleet composition
		% Electric vehicles
		Achieved and expected GHG reductions
		Net revenues and % high-emission industries
		Climate change mitigation
	E2 – Pollution	Pollution of air, water, and soil
		Microplastics generated and used
		Total emissions of pollutants
		Substances of concern generated or used
		Financial effects of risks and opportunities
		Revenue related to concerning substances
		Provisions for environmental protection
		Provisions for remediation costs
	E3 – Water and marine resources	Total water consumption
		Total water recycled and reused
		Total water stored (incl. changes)
		Water intensity ratio
		Financial effects of risks and opportunities
	E4 – Biodiversity and ecosystems	Financing effects of biodiversity offsets
		Site number and area around protected areas
		Total use of land and nature-oriented area
		Financial effects of risks and opportunities
	E5 – Resource use and circular economy	Total weight of products and materials
		% Biological materials & recycled components
		Recyclable content in products and packaging
		Total waste generated
		Financial effects of risks and opportunities

Reporting challenges

The reporting challenges are divided in three themes. Under the first theme of expectations and understanding, De Vries found that companies struggle to comprehend CSRD's impact on their organization and assess their current reporting maturity [6]. Understanding the structure and guidelines together with the establishment of new processes, requires an investment of time and resources (either through in-house expertise or external consultants), which poses to be troublesome for several organizations [6], [7], [12], [51], [52]. Furthermore, forming multidisciplinary teams is challenging, and the CSRD demands significant organizational changes beyond new performance metrics [53].

KPMG found that most of the 200 companies they examined fell short of the CSRD requirements [33]. Additionally, some organizations are unaware of the directive's imminent impact and additional workload [52]. Companies must change from traditional,

analogue reporting to more automation for an active and continuous approach, due to CSRD's traceability and data handling requirements [12], [54]. The significant data volume requires this change. Additionally, the CSRD targets a group of companies unaccustomed to non-financial reporting, and many lack the necessary structures [7], [47]. Finally, organizations struggle with the materiality assessments, facing challenges in identifying the relevant ESRS areas and determining the required level of detail for reports [55].

Considering the software theme, PwC found that 64% of the companies in their research face challenges with the CSRD's complex technical implementation, and a similar number struggle with their data base. Nearly one-third are unsure about suitable software solutions (as of mid-2023). More than a quarter plan to use Excel, while one-fifth aim to adopt dedicated sustainability software [13]. Around the same time, LeanIX found 62% using Excel spreadsheets to track data [11]. However, both Spek & de Vries and Markendahl noted that spreadsheets and manual data collection will become insufficient due to increased data amounts [30], [55]; while some companies are still performing work on paper.

Under the theme of data challenges, 61% of organizations questioned by PwC already collect KPIs relevant to CSRD, but over half expressed that scarce resources pose a challenge [13]. Here, the value chain is the most prominent obstacle for CSRD compliance: sustainability performance and reporting maturity are in a way dependent on an organization's value chain [12], [14], [30], [54], [55], [56]. This can be especially complicated with value chains reaching outside of Europe and data related to Scope 3 [52], [55].

While an abundance of data is available, accessibility and retrieving it from suppliers are major issues, stemming from a lack of knowledge and resources [12], [56]. Different vocabularies among stakeholders complicate data aggregation and extracting conclusions [57], while data reliability and quality can be questionable when it is being collected from distant suppliers [55].

Several companies expressed concerns about data availability and credibility, especially in ensuring reliability and validation of measures; ESG data sources often lack the completeness, reliability and transparency of financial information [6], [56]. Moreover, integrating data from isolated sources and departments and general data scarcity were mentioned [6]. Obtaining, preparing and presenting the necessary data is a notable challenge, especially as firms are assessing what (unstructured) data is available and what needs to be gathered [47], [52], [55]. De Vries' research shows that mainly E2 – Pollution, E4 – Biodiversity and S4 – Consumers and End-users are currently barely reported on.

An important consideration is the combination of quantitative and qualitative data. Quantitative data must be tagged according to the digital taxonomy and accompanied by a significant number of qualitative explanations. Together, the reporting should clearly outline the effects, risks, and opportunities each standard brings to the organization and its value chain [2], [56].

De Vries found data integration to be the main challenge for organizations, with software connectivity issues causing hesitation to adopt new software solutions [6]. This is confirmed by Matilla & Sasi, who also identified "*getting all pieces into a system in order to report*" as the true challenge [52, p. 28].

Motivation to report

Looking at motivation, Bauer & Greiling found Austrian Non-Profit Organizations (NPOs) lack motivation in environmental management. While they have begun integrating environmental considerations, their focus is on long-term planning rather than short-term operational practices. Organizations set to start in 2025 are still in the early stages of their journey. Above all, environmental matters are prioritized lower than social and financial concerns [58].

External pressures from authorities, investors, the market, and society can drive motivation. Transparency and effective communication are necessary for effective sustainability reporting. However, Eklund and Vaaler's research indicates that many companies are mainly motivated out of obligation: "*neither we nor other companies would report as extensively on as many topics if it were not for these regulations*" [54, p. 53], reflecting a "*tick-the-box*" approach brought up by Pizzi et al. [9].

The research by Glaveli et al. supports EFRAG's statement that "*sustainability is perceived as a threat to the business, rather than as an opportunity for development and innovation*" [59, p. 11]. Nonetheless, motivated companies with mature sustainability disclosures aspire to stand out, gain a competitive advantage, and build trust with their stakeholders.

Digitalization in a CSRD context

Considering digitalization, Svensson explored challenges and opportunities resulting from the new CSRD. Companies have yet to focus on digital transformation for sustainability reporting, despite the vast amounts of data making the use of traditional methods impossible [12]. Apart from numerous challenges (see from page 35), Svensson identified that accurate reporting can be a competitive advantage. Furthermore, increased automation in data gathering and control enhances risk awareness, allowing better prevention or mitigation [12].

Atanasov discusses digital tagging and the challenges of organizing, accessing, evaluating, and acting on relevant data [2]. Companies that master digital transformation will find significant opportunities, with improved economic results linked to digitalization and sustainability capabilities. Digital transformation enhances sustainability reporting, showcasing a strong connection between the two [2].

Eklund & Vaaler investigated the transition to CSRD and ESRS, focusing on challenges for Norwegian companies regarding CSRD requirements and digitalization [54]. The well-known main challenge is obtaining information throughout the value chain, followed by the need for significant improvements to IT systems to fulfil reporting requirements. Furthermore, concerns include manual procedures and developing interactive systems,

with reports potentially becoming overly comprehensive, leading to a lack of oversight and significance [54].

Pizzi et al. emphasize digitalization opportunities, noting that sustainability standardization fosters new reporting tools and technologies [9]. Moreover, digital transformation can aid the transition of European companies to become more sophisticated in accountability approaches.

De Vries identified challenges related to digitalization (from page 35), stemming from integration issues and hesitation towards new software solutions [6]. Vărzaru focuses on opportunities, examining the impact of new technologies on sustainability reporting [10]. Digital technologies are said to improve the effectiveness of sustainability accounting and reporting tools, enabling better tracking of social and environmental activities.

3.1.2 Model-based analysis techniques

Considering the main challenge of data consolidation, solution areas must be investigated. The second research question, "*What model-based analysis techniques can be applied in the context of CSRD?*", is therefore concerned with the discovery of architectural frameworks and modelling languages, and how they can be applied in a CSRD context. Thus, first the concept of Enterprise Architecture is explored. Then, architecture frameworks, modelling methods, and languages are investigated and assessed based on their suitability for CSRD.

Enterprise Architecture

Enterprise Architecture (EA) is a well-known starting point for model-based analysis techniques in an organizational context. It is a strategic discipline, playing a pivotal role in the development and execution of an enterprise software strategy through company-wide integration [60], [61]. Since the late 1980s, EA has evolved as a discipline and method to manage information systems, associated business elements, and complexity within an organization [62], [63]. The EA Body of Knowledge defines it as:

> *"a practice, which analyses areas of common activity within or between organizations, where information and other resources are exchanged to guide future states from an integrated viewpoint of strategy, business and technology"* [64, p. 81].

In other terms, Enterprise Architecture Management (EAM) is a field that outlines the integral role of IT within an organization. It focuses on the interaction among various elements within the IT and business operations of the enterprise [65].

There are several principles, frameworks, methods, and models used to design and bridge the communication gap between architects and stakeholders. EA achieves this with the visualisation of business and organization from a bird's-eye view [61], [64], [66]. Furthermore, it can help organizations optimize business activities, improve strategies, and better utilize IT [67].

EA modelling aims to clarify enterprise strategies, visualize business processes, and model information systems. This aids resource management, organizational improvement, information strategy adjustment, and new business value creation [66]. EA modelling is crucial for IT infrastructure analysis, but models can become large and complex when multiple data sources cover several enterprise domains. In this case, EA model creation can even be automated to a certain extent [63].

Besides the bird's-eye view, EA focuses on software architecture (SA), defined as *"the structure or structures of the system, which comprise software components, the externally visible properties of those components, and the relationships among them"* [68, p. 639]. SA considers both the system's internal and external environments, though architectural knowledge is often undocumented and can easily get lost [69]. Furthermore, EA supports development of the IT landscape, which is architecture design on a large scale beyond a single system [70]. Lastly, the ability to evolve software to meet changing requirements (software evolution) can be realized through an architectural solution [71].

Architecture frameworks

The core concept within EA is the architecture framework, which aligns various resources (such as IT) with the organization's present and future needs [67]. Frameworks like the Zachman Framework and TOGAF can be used in synergy; Zachman provides an EA ontology in a matrix approach, while TOGAF offers a concrete EA method in a layered approach [60], [72]. The EA methodologies guide the implementation of these frameworks and outline how organizations transform from their as-is states to to-be. TOGAF, for instance, provides detailed guidance, whereas Zachman consists of general guidelines [67].

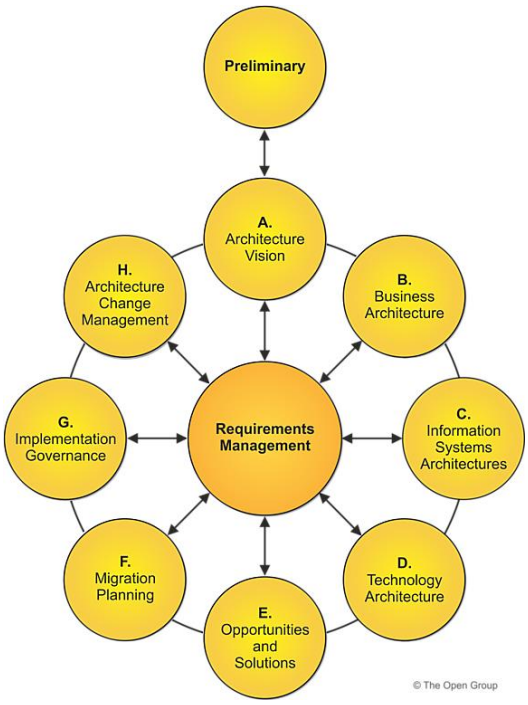


Figure 9 - Architecture Development Method by The Open Group (n.d.)

To get a better understanding of such frameworks, TOGAF is discussed in more detail. The core of the framework is the Architecture Development Method (ADM) which describes a method to develop and manage an Enterprise Architecture's lifecycle. A complete EA must cover four architecture domains, defined as business, data, application, and technology [73], [74]. The business architecture is concerned with the elements describing the business design; the data architecture is about storage and accessibility of sources; the application architecture looks at the design and interaction of applications; and the technical architecture is concerned with hardware, software, and their interactions [73]. The ADM is depicted in Figure 9.

Modelling methods

"*EA Modelling refers to the creation and management of architectural models*" [62, p. 76]. It is a significant element to depict the cross-sectional aspects and relations within an organization. Moreover, EA modelling offers perspectives through different viewpoints for strategic IT alignment, guiding an organization toward its desired future state [67]. The integration of these perspectives provides a thorough and consistent description of the organization's state [75].

A modelling method consists of a modelling language; modelling technique and procedure; and modelling algorithms and mechanisms [72], [75]. Furthermore, we can speak of modelling tools, modelling concepts, and modelling deliverables [62]. The modelling language contains the syntax, semantics, and notion. Then the procedure is the description of specific steps that must be followed to create valid models. Furthermore, the algorithms are executed on the modelling language either through a generic[8] or specialized algorithm[9] [75].

Zooming in on modelling languages, Lankhorst argues that they must focus on inter-domain relations. Furthermore, there should be a formal foundation that guarantees model interpretations are clear and that they can be automatically analysed. Moreover, there must be enough flexibility to customize model visualization, such that they can meet the unique information requirements of various stakeholders [61].

ArchiMate, a modelling language founded by The Open Group (also the creators of TOGAF), is a popular tool for designing EA. ArchiMate is one of the most comprehensive modelling languages in the EA domain and favours a concise language design. Furthermore, it provides complete support for the TOGAF framework [64], [76]. Originally, the language has been developed "*to provide a uniform representation for diagrams that describe enterprise architectures*" [77]. The idea is that the architectures can be represented over time, including transformation and migration.

The language is organized into three layers (business, application, and technology) and two extensions (motivation and implementation and migration). The business layer is concerned with business process and architecture elements; the application layer describes software applications; and the technology focuses on hardware and software infrastructure [60], [73], [75]. The elements in the language are grouped into active, behavioural, and passive structures. Furthermore, a meta-model and graphical notation are provided to visually describe and analyse the three EA domains (the layers), their relationships, and dependencies [60].

[8] An example of a generic algorithm is a shortest-path algorithm.

[9] Specialized algorithms incorporate process and capacity simulations.

Research by Zhou et al. concluded that compared and contrasted with languages such as UML, BPMN, BMM and SoaML, ArchiMate is the most potent modelling notation when compared for visual expressiveness. Furthermore, the language is most commonly utilized by scholars [66].

Table 4 portrays an overview of each of the languages and the previously defined criteria (Sub-section 1.3.3). These investigate whether the language can model from an IT perspective; what the level of granularity is for the language; whether it can describe relations between data sources; and whether the language is diversified or if it has a focused purpose. Based on the scores, ArchiMate and UML are most suitable, modelling from an IT perspective and having the ability to model relationships. Nonetheless, ArchiMate then still has the advantage of having a low granularity and not being focused on one specific purpose, having the ability to model the IT architecture from a high-level. This is especially valuable in a CSRD context, where it is important to look at interconnections and dependencies.

Table 4 - Assessment of modelling languages

Language	Sources	IT perspective	Granularity	Relations	Diversified
UML	[60], [61]	x	High	x	
BPMN	[60], [61]		Medium		x
ArchiMate	[60], [66], [73], [75], [64], [76], [77]	x	Low	x	x
BMM	[60], [72]		Medium		
SoaML	[60], [78]	x	High		
FAML	[60], [73]	x	High		

3.1.3 Data consolidation techniques

Moving from techniques that can help visualize potential integration pathways, this sub-section focuses on techniques that ensure these integrations, allowing for data consolidation. This is necessary to bring data together for CSRD reporting. To gather information on data consolidation techniques, first, an exploratory literature search was. In doing so, four techniques were identified and included in this research. Furthermore, a systematic search was performed to find more details on data warehouses and data lakes. Altogether, these results are portrayed in the following sections to form a base for RQ 3: "*What are best practices for consolidating data from multiple sources for reporting purposes?*".

Data consolidation

According to Powell and Smalley, "*data consolidation means bringing together data from various sources and assembling it within a single location*" [79]. It serves as a single point of access for users and allows the generation of data insights while dealing with a myriad of data formats. Several benefits of data consolidation are improved decision-making; cost reduction through efficiency; time savings due to a Single Source of Truth (SSoT); and smoother emergency operations related to disaster recovery [79].

Especially with the growing amount of business data (big data), consolidation and sharing has become an important service among organizations [80], [81]. The integrated data allows for quick information presentation, in an accurate and real-time manner. This supports the decision-making processes, predictive analysis, and company policies within an organization [82], [83]. Furthermore, consolidation can aid in gathering useful and transferable information to help highlight problems, especially with input coming from various agencies [81].

Extract-Transform-Load

One of the most well-known and important techniques for data consolidation is ETL (Extract-Transform-Load). First, information is extracted from multiple data sources using ETL tools. These tools can automate workflows and help with understanding and using data in the desired location [81]. Then, the data is automatically transformed into a conventional informational format, according to certain rules and techniques. This process includes cleaning, filtering, joining, and validating data [84]. As a final step, data is loaded into a designated, centralized destination [79], [80], [85], [86].

Several researchers identified some challenges with the ETL approach, including difficult access to critical data; long processing times; and resulting data not meeting standard quality. Furthermore, there are some shortcomings related to cost, time, performance, process, and continuous improvement [80], [85]. The reason for many of these is that information is not available during the ETL process, resulting in increased job runtime and costs [82]. Additionally, with the growing data volumes data is becoming increasingly complex. Therefore, it is also more difficult to develop effective ETL solutions [85].

Extract-Load-Transform

Because of the aforementioned challenges with ETL, a novel technique was created: ELT (Extract-Load-Transform). In this process, data is again first extracted, but the second step is to load it into the warehouse. The data stays in this space and is analysed from various perspectives in the organization, ultimately leading to its transformation [79], [80]. Compared to ETL, this new process saves costs and extra processing, while further reducing network congestion. Additionally, the ELT technique serves higher performance and scalability, making the management of infrastructure integration easier. Lastly, ELT works better for larger data volumes, which is why it is often considered as the superior method [82].

Data warehouse

A data warehouse is a centralized repository, which consolidates data from various sources using ETL tools. This central data collection allows for a higher degree of data security than in a decentralized landscape [79]. For organizations, data warehouses can provide strategic information to support decision-making, consolidating operational data sources from several business units [82]. Furthermore, they provide an integrated and consistent view of the consolidated data [83].

A data warehouse functions as a repository for structured, filtered, and processed data which has a specific purpose. One of the most well-known definitions is by Inmon, who defined a data warehouse as *"a subject-oriented, non-volatile, integrated, time-variant collection of data in support of management decisions"* [87, p. 3], [88], [89]. The warehouse serves as the primary origin of data for reporting and analytical purposes, employing online analytical processing (OLAP) [87], [89].

When using a data warehouse, the data structure in the database source must be known, as data needs to be transformed according to the destination structure. Furthermore, transferred data should meet the data types expected by the warehouse [90]. This is because data warehouses work according to a fixed schema [91]. The architecture of a data warehouse is visualized in Figure 10, adapted based on Harby & Zulkernine [88], Kutay [92], and Nambiar & Mundra [87].

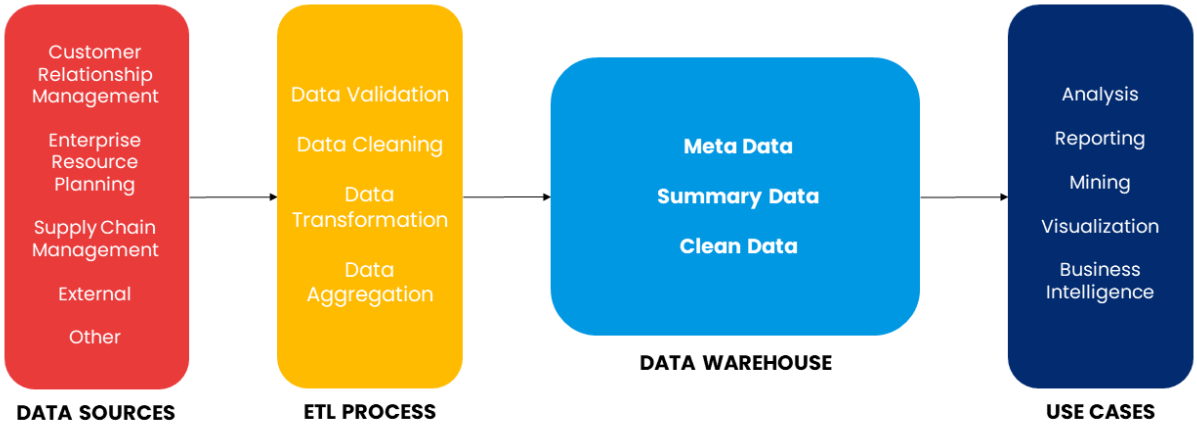


Figure 10 - Visualization of the data warehouse architecture

Data warehouses are used both on-premises and in the cloud, storing enterprise data and supporting business intelligence (BI) and analytics applications [87], [92], [93]. The modern cloud data warehouse does provide some advantages over the traditional warehouse, as it supports all types of data (semi-structured and unstructured), provides faster response time, cost savings, easier scaling up or down, and increased flexibility [89], [94].

Benefits of a data warehouse are consistent performance, ease of use, a Single Source of Truth, fast response, efficient decision-making, standardization, data integration, and cross-functional analysis [90], [92], [93], [94]. Moreover, the historical data stored in the warehouse can be used to predict the future using machine learning algorithms, allowing organizations to make data-driven decisions [94].

The most widely used architecture for data warehouses is the three-tier architecture. Within the bottom tier, data undergoes cleansing, transformation, and loading processes using backend tools. Then the middle tier functions as an OLAP server, offering a conceptual view of the database. The top tier, known as the front-end client layer, includes tools and an API

used for connecting to and extracting data from the data warehouse. These include query tools, reporting tools, analysis tools, managed query tools, and data mining tools [87].

Nambiar & Mundra suggest using ETL with expensive in-house analytics systems and ELT with cloud data warehouses [87]. Using ETL in combination with a data warehouse can result in loss of information while trying to get data into the fixed scheme [91]. Shiyal adds that one has to wait for the ETL process to be finished, whereas ELT offers faster access in cloud environments [94]. Furthermore, ETL is limited in handling high-speed data ingestion and variations in the structure of incoming data [88].

Data lake

Proposed in 2010, data lakes enable cost-effective storage and flexible scalability by preserving data in its raw form to overcome data warehouse limitations [79], [87], [91], [94], [95], supporting high-speed and unstructured data storage [88], [93]. Unlike warehouses, which process and package data, data lakes primarily serve as repositories, storing various data types without alteration [87]. This makes them suitable for storing large volumes of data that are easily retrievable and maintainable [94]. Data lakes often complement data warehouses in data processing tasks [87].

Data lakes store unstructured, semi-structured, and structured data at a large scale. Similarly to warehouses, lakes support business decisions via big data analysis, machine learning, and real-time analytics, but lakes offer greater flexibility and scalability due to the lack of a strict schema [87], [90], [92], [96], [97], [98]. Data transformation rules are applied only upon retrieval. Rich metadata helps discover patterns in the raw data and understand the stored datasets in the data lake [99], [100]. Without metadata and data governance processes, there are no insights and the lake can turn into a "data swamp", wasting resources and value [88], [101]. Figure 11 illustrates the data lake architecture, adapted based on Harby & Zulkernine [88], Kutay [92], and Nambiar & Mundra [87].

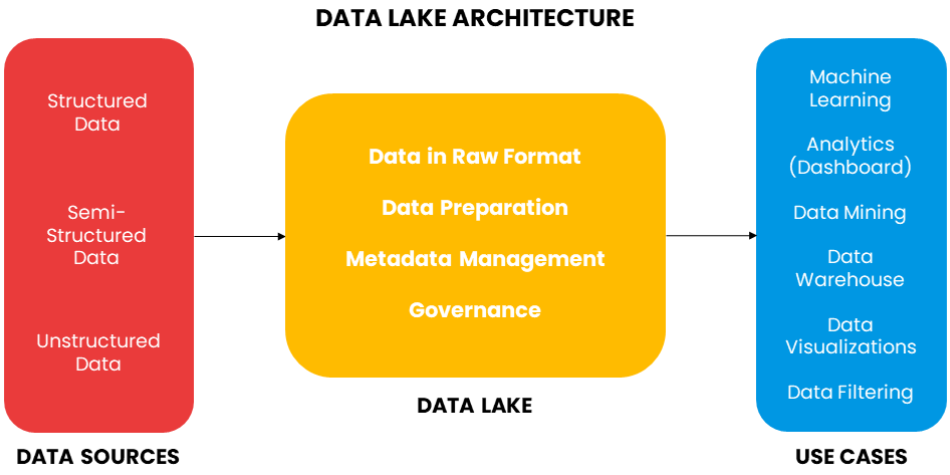


Figure 11 - Visualization of the data lake architecture

The various benefits of data lakes therefore include greater scalability; easier storage and accessibility of different data types; low cost storage and processing; support of advanced analytics and data science techniques; and programming support [90], [92], [93], [98].

Consequently, data lakes enhance the capture, refinement, archiving, and exploration of raw data [97]. Data scientists benefit from reduced time and resource demands for data preparation compared to the process of data warehouses [92], [101].

While data lakes are generally cost-effective, Nambiar & Mundra highlighted they can be expensive to implement and maintain, requiring domain experts and engineers to manage them. There is, however, a shortage of data science professionals [87]. Furthermore, capturing large amounts of data from various sources raises questions about data quality, which may vary by use case [96], [102], affecting data reliability and security [92].

There is no common consensus on data lake architecture design, apart from it involving multiple layers (data ingestion, maintenance, and exploration) [100]. However, there are two well-known architectures for data lakes. The Zone Architecture divides a lake into zones for organizing data types, automating tasks, and restricting direct access to raw data, which increases administrative efforts. The Lambda Architecture reduces complexity, simplifying the creation of production systems with established workflows while keeping all raw data sets for future use. The downside of this architecture is that it is more rigid [91].

Data lakes employ ELT, where as much data is gathered as possible, and transformation occurs based on specific use cases when data consumers build models [90]. ETL is considered insufficient for this purpose [99].

Data lakehouse

From the research on data warehouses and data lakes, it became apparent that nowadays there is a new technique called a data lakehouse. As described by Armbrust et al. in 2021, they emerged due to the common scenario where raw data from a data lake is utilized as input for an ETL process to fill a data warehouse [91].

The lakehouse combines the large-scale, low-cost, flexible raw data storage of a data lake with the analytics capabilities (transaction support, indexing, caching, and metadata management), governance standards, and data quality of a data warehouse [87], [99], [101], [102], [103].

Defined by Armbrust et al., a data lakehouse is "*a data management system based on low-cost and directly-accessible storage that also provides traditional analytical DBMS management and performance features such as ACID*[10] *transactions, data versioning, auditing, indexing, caching, and query optimization*" [93, p. 38], [104, p. 3]. Figure 12 visualizes a data lakehouse architecture, adapted from Oreščanin & Hlupić [101] and Harby & Zulkernine [88]. An in-depth explanation of the different layers is included in [101].

[10] Atomicity, Consistency, Isolation, and Durability

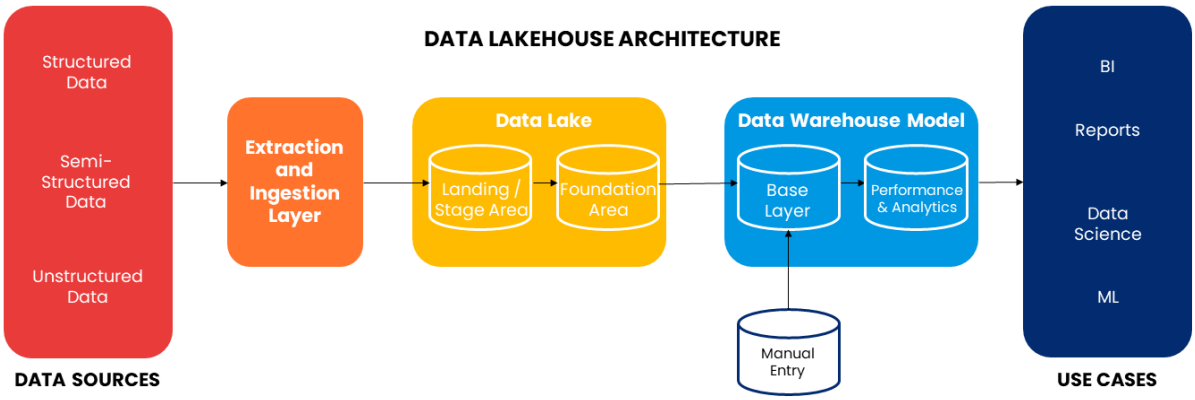


Figure 12 - Visualization of a data lakehouse architecture

The lakehouse supports storage and analysis of text, images, audio, video, etc., while deriving insights from unstructured data through a valuable metadata layer on top of an object store. This setup allows for independent scaling of storage and computing resources, saving whenever computing resources are not used [91], [94], [102], [103].

Unlike modern data warehouses, lakehouses do not require copying data to a relational database; data can be stored in various zones, accommodating both historical and real-time data [90], [99]. They also support ACID transactions, ensuring data consistency and integrity [87], [89], [90], [99], and enable big data analytics, AI, streaming capabilities, machine learning tools, and OLAP queries for business intelligence [84], [88], [97].

Compared to modern data warehouses, lakehouses do not require copying data from the lake to a relational database; data can be stored in various zones, accommodating both historical and real-time data [94], [103]. Furthermore, lakehouses support ACID transactions, ensuring data consistency and integrity [91], [93], [94], [103]. Additionally, lakehouses enable big data analytics, artificial intelligence, streaming capabilities, machine learning tools, and OLAP queries for business intelligence [88], [92], [101].

Challenges include the monolithic architecture, the infancy of the technique, a lack of products and options, and scarcity of skilled technical resources [94]. Considering the novelty of the lakehouse, it remains uncertain whether it can fulfil its promises [92]. Regarding ETL vs. ELT, Mazumdar et al. explain that data is again first extracted and loaded, followed by necessary transformations [103].

3.2 DISCUSSION OF LITERATURE FINDINGS

Aiming to answer, "*What methods and techniques can be utilized to integrate various data sources within companies' IT architecture and along their value chain, facilitating data consolidation for CSRD reporting*?", the first section of this chapter focused on three research questions. Therefore, these will now be answered in this discussion.

3.2.1 CSRD reporting

Looking back, this chapter first investigated RQ 1, "*What does CSRD reporting entail?*". Initially, large organizations must comply, followed by other companies meeting specific criteria. Furthermore, the ESRS cover environmental, social, and governance-related topics

in 10 topical standards. Reporting on these topics must cover a companies' entire value chain (including Scope 3 emissions). Sector-specific standards, postponed until mid-2026, will also apply to companies in certain industries. Additionally, materiality is a key component of the CSRD, requiring companies to assess both financial and impact aspects.

CSRD reporting requires digital tagging, with companies submitting reports in the European Single Electronic Format, using XHTML in combination with XBRL tags. Moreover, an external audit is mandatory to assure sustainability information matches financial information quality. However, organizations are not obliged to share sensitive information if the overall relevance of disclosure is maintained. An overview of relevant KPIs is provided on page 33.

Companies face three main categories of challenges with CSRD reporting. First, there are challenges regarding expectations and understanding the requirements. Companies struggle to comprehend the CSRD, including structure and guidelines, which simultaneously requires an investment of time and resources. The CSRD demands a change in processes, where many organizations are currently falling short of its requirements. Traditional reporting ways must be changed, while at the same time a materiality assessment must be conducted.

Second, software challenges arise as traditional spreadsheets and manual data collection become inadequate, and companies are uncertain about suitable software. Third, data challenges include reporting on the entire value chain, data availability and quality, retrieving data from suppliers, and integrating isolated data sources. Furthermore, qualitative and quantitative data must be combined, and quantitative must be tagged according to the digital taxonomy. Lastly, the main challenge is said to be software connectivity and integration: how to get all pieces into a system for reporting purposes.

3.2.2 Model-based analysis techniques

Moving from this main challenge to a potential solution, this chapter explored model-based analysis techniques. Answering RQ 2, "*What model-based analysis techniques can be applied in the context of CSRD?*", Enterprise Architecture is a valuable starting point. EA helps visualize an organization's business and IT landscape, providing stakeholders with a comprehensive overview. Using frameworks as implementation guides and modelling languages for visualizations, EA clarifies relationships between processes and applications. This clarity serves as a foundational element for CSRD reporting, helping to map relationships and visualize data flows, such that organizations understand the origins of data, connections between applications, and potential integration pathways. This visual overview facilitates understanding current operations and informs strategic decisions for aligning business processes with sustainability goals.

This research highlights that the methods created by The Open Group are the most popular model-based analysis techniques. Both TOGAF and ArchiMate are most frequently mentioned in literature, with ArchiMate being *the* visually expressive modelling language [66]. All in all, we expect that ArchiMate is best applicable in the context of CSRD. Its popularity in both the academic and professional world [58] proves its expressiveness and

use, offering a comprehensive view of the CSRD landscape from business, application, and technology perspectives.

3.2.3 Data consolidation techniques

Utilizing a modelling language, several to-be landscapes can be modelled that include a data consolidation technique. These techniques are useful in a CSRD reporting context because they focus on data integration. By consolidating data from various sources into a central Single Source of Truth, organizations establish a reliable foundation for their reporting processes. Due to the magnitude of data, this is necessary to create the CSRD reports. This centralized data repository ensures consistency and accuracy across reports, enabling stakeholders to access a unified and trustworthy version of data. Subsequently, this streamlined approach simplifies the extraction of reports.

Investigating the third research question, "*What are best practices for consolidating data from multiple sources for reporting purposes?",* this chapter discussed ETL, ELT, data warehouses, and data lakes. Investigating the details of data warehouses and data lakes even further with a systematic literature search, novel findings included the data lakehouse. With many papers including longer in-depth discussions of each technique, this chapter aimed to provide a comprehensive overview of each of them.

Several researchers discuss the direct differences between data warehouses, data lakes, and sometimes even the new data lakehouse. Based on the tables by Nambiar & Mundra [87], El Aissi et al. [90], Hukkeri et al. [97], Harby & Zulkernine [88], Janssen [93], and Kutay [92] and additional information by Zouari et al. [96] and Chen [100], Table 5 is constructed to provide an all-encompassing comparison of the three techniques. Figure 13, created by Lorica et al. [102], shows a simpler overview. Some fields for the data lakehouse have been left empty in Table 5, as no distinct answer was found in the synthesized literature.

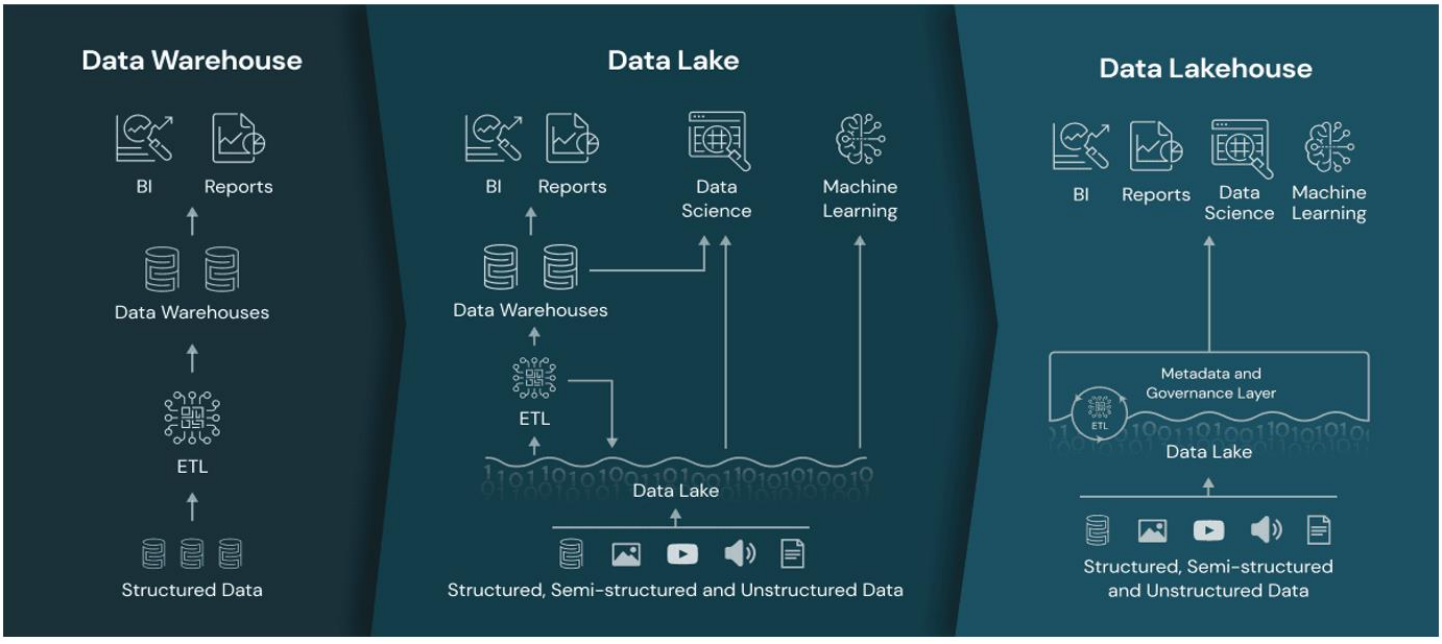


Figure 13 - Visual comparison of the Data Warehouse, Data Lake, and Data Lakehouse by Lorica et al. (2020)

The main differences can be identified as the structures of data, the way data is organized, and the form of analysis. Furthermore, a key difference is the purpose of data warehouses which is very specific compared to the undefined usage of data lakes [87]. Together, all insights can be used for the discussion on which technique is most appropriate in the CSRD reporting context.

Table 5 - Comparison of the Data Warehouse, Data Lake, and Data Lakehouse

Parameter	**Data Warehouse**	**Data Lake**	**Data Lakehouse**
Data	Focus only on business processes	Store everything	Story everything, with possibility to focus
Processing	Highly processed data	Data are mainly unprocessed	Combination of both
Type of Data	Mostly in tabular form and structure (clean)	Unstructured, semi-structured, sometimes structured (raw)	Unstructured, semi-structured, or structured (raw)
Schema Definition	Before data collection – ETL (schema-on-write)	After data collection – ELT (schema-on-read)	Both schema-on-write and schema-on-read
Task	Optimized for data retrieval; data analytics and BI	Share data stewardship; efficient in developing real-time applications; ML and AI	Suitable for all use cases (data analytics, BI, ML, and AI)
Agility	Less agile and has a rigid configuration	Highly agile and flexible, can configure and reconfigure as needed	Highly agile and flexible, can configure and reconfigure as needed
Use cases	Widely used by business professionals and business analysts	Used by data scientists, data developers, and business analysts	Wide use case by combining the structure and simplicity of a DW with the wider use cases of a DL
Ease of use	Fixed schema makes data easy to locate, access, and query data; easy to report	Time and effort are required to organize and prepare data; analysing vast amounts of data without proper tools can be difficult	Simple – interfaces are provided that are similar to DWs
Storage and Costs	Large data volumes at a moderate cost; difficult scaling	Low-cost storage of extreme data volumes; fast and flexible	Low-cost storage of extreme data volumes; fast and flexible
Architecture Design	Hierarchical (with folders and files)	Flat (each data element has its ID tag)	Monolithic
ACID Conformity	Data is recorded in an ACID-compliant way to ensure integrity	Non-ACID compliance: updates and deletes are difficult procedures	Ensures consistency through ACID-compliance
Security	Allows better control of the data	Offers less control	
Data Processing	Time-consuming to introduce new content	Helps with fast ingestion of new data	Helps with fast ingestion of new data
Data Granularity	Data at the summary or aggregated level of detail	Data at a low level of detail or granularity	
Tools	Mostly commercial tools	Can use open-source tools	
Limitation	Complex joins	Complex processing; might lead to data swamps	Infancy of the technique and scarcity of skilled technical resources

Considering all data consolidation techniques, the most well-known is ETL, nowadays having a predecessor called ELT. With different advantages and disadvantages, these techniques cannot be used standalone; there is always a place where the data needs to be loaded.

Comparing data warehouses, data lakes, and data lakehouses, each has its advantages for data consolidation (Table 5 and Figure 13). Taking the differences outlined in the table and figure into account, it is likely that data warehouses are more suitable for smaller organizations; data lakes for larger organizations with more resources; and the data lakehouse for (larger) organizations that would like to work with novel techniques, who have the expertise in-house or who can hire experts.

Focusing in-depth on data warehouses, they are designed for reporting and business intelligence and are user-friendly without needing data science expertise. Considering costs, data warehouses are said to be expensive to set up and maintain, but cloud data warehouses do not require the purchase of physical hardware. Therefore, they are quicker and cheaper to set up and scale. Smaller companies with a more compressed value chain might be able to work with structured data, where a data warehouse can easily consolidate data.

Data lakes, suitable for larger organizations, store both structured and unstructured data, beneficial for extensive value chains. The scalability and flexibility of the data lake support diverse use cases like financial predictions but require proper governance and data science expertise to avoid becoming a data swamp. The data lake does provide the possibility to move beyond reporting and business intelligence, incorporating machine learning and artificial intelligence.

The last – and most novel – technique is the data lakehouse, combining best of both data warehouse and lake. Whenever organizations have the resources and ambition to work with this innovation, it provides several advantages: storage of structured and unstructured data, simple usage through data warehouse-like interfaces, suitability for all use cases, and low-cost storage. However, they are still new and require skilled technical resources, posing a challenge for organizations. Nonetheless, if an organization is willing to take this step, it can give them great benefits in facilitating data consolidation for CSRD reporting.

In summary, best practices include using a data warehouse with ETL (or ELT), a data lake with ELT, or a data lakehouse with a combination of both. However, all these options do depend on each organization's specific situation and ambitions. Regardless of the exact technique, they are all capable of integrating data and information from various resources into a central space for CSRD reporting purposes.

3.3 METHODS AND TECHNIQUES TO INTEGRATE VARIOUS DATA SOURCES

Finally, the answer to the main research question of this chapter consolidates all relevant aspects into a comprehensive overview:

> "*What methods and techniques can be utilized to integrate various data sources within companies' IT architecture and along their value chain, facilitating data consolidation for CSRD reporting?*"

Considering the methods, Enterprise Architecture would be a good starting point to manage the IT landscape in a CSRD context. It can help analyse both organizations' own landscape and how the landscape relates to their value chain. EA modelling, focused on clarifying strategies, visualizing business processes, and modelling information systems, allows for depiction of cross-sectional aspects and relations within and outside of an organization.

Providing a concrete answer on the side of IT landscape modelling, based on findings in literature, ArchiMate is found to be a comprehensive language to visualize the CSRD landscape. Considering that this language is often paired with TOGAF as architectural framework, the Architecture Development Method is a good backbone for development of the models. Furthermore, ArchiMate could be further enhanced with UML whenever necessary to go more into technical details.

For a data consolidation technique, the decision depends on the context and situation of each organization. However, it has been defined that either a data warehouse, data lake, or data lakehouse can bring the necessary capabilities to consolidate data from various data sources. Possible scenarios for the application of these techniques have been described in Sub-section 3.2.3, but all can be suitable in the context of CSRD reporting.

4 EMPIRICAL PROBLEM INVESTIGATION

Additional to the literature review that aimed to find both challenges regarding CSRD reporting and potential solutions for these challenges, empirical research is conducted. The objective for this part of the study is to find the practical challenges and experiences, including best practices in a CSRD context. Based on the methodology as described in Section 2.3, this chapter starts with the interview procedure (Section 4.1), followed by the method for data analysis (Section 4.2). Subsequently, the results are presented in Section 4.3, followed by a discussion to answer the sub-research questions in Section 4.4 and concluding remarks on RQ-E in Section 4.5.

4.1 INTERVIEW PROCEDURE

To ensure a comprehensive understanding of the problem in the current situation, interviews are conducted to find additional insights to the previously conducted literature review. The aim of these insights is to answer RQ-E:

> *"What are the practical challenges and experiences that companies encounter while implementing CSRD reporting, particularly in terms of motivational drivers and technical boundaries?"*

As previously stated under the scope, the study encompasses organizations that undergo the process of implementing the CSRD, either as a result of intrinsic motivation or due to their classification within the first or second mandatory group (large & listed organizations and large organizations).

In order to reach out to these organizations, the researcher initially contacted customers and leads of Flawless Workflow, that were known to have initiated the CSRD implementation process. Furthermore, well-known organizations that were sufficiently large to be included in the mandatory group were also contacted. Finally, organizations for which it is publicly known that they intrinsically want to comply with the CSRD were sought out.

Contact was made with specific employees involved with the CSRD, or they were located through their colleagues. A total of eight organizations expressed their willingness to participate in an interview. Table 6 presents the information pertaining to the interviews conducted. This includes the tags for the interview and the company, the operating sector, the position of the interviewee within the company and whether the organization is part of a group for which it is mandatory to create a CSRD report.

Table 6 - Demographics of organizations and interviewees; IP = In-person interview, O = Online interview via Teams

Interviewee Tag	Company Tag	Sector	Position	Group	IP/O
INT1	Company A	Retail	Manager Inbound Logistics	Mandatory	IP
INT2	Company B	Retail & metal processing	Innovation Manager	Mandatory	O
INT3	Company C	Machine production	Intern	Mandatory	IP
INT4	Company D	Education and science	Advisor Policy & Analytics	Intrinsic	IP
INT5	Company E	Oil and gas	Sustainable Finance Project Manager	Mandatory	O
INT6	Company F	Social housing	Financial Controller	Intrinsic	O
INT7	Company G	Logistics	Strategic Program Manager CSRD	Mandatory	O
INT8	Company H	Equipment manufacturing	Head of Finance	Mandatory	IP

The interview questions are included in Appendix 9.6. It should be noted that there were slight differences in the questions asked and their order for each interview, depending on the flow of the conversation with the interviewee. Prior to the interviews, certain keywords were identified for potential answer directions for some of the questions. These could be used as examples if the interviewee found it difficult to understand the direction of a particular question.

Depending on the interviewee's preference, interviews were conducted either online or in person. For each interview, the interviewee was sent an information sheet beforehand and was asked to sign an informed consent form. Based on their consent, the interview was audio-recorded in person or video-recorded via Microsoft Teams. Additionally, written notes were made. This is all done according to the ethics approval by the BMS ethics committee of the University of Twente. One interview took between 45 and 60 minutes, starting with an introduction about the researcher and the research.

4.2 DATA ANALYSIS

For the interviews via Microsoft Teams, the transcribing feature was used to provide a basis for the transcription. For the in-person interviews, the transcription function of Word was used to transcribe the audio file. Each transcript was then checked and corrected by listening to the audio or video recording. In addition, identifying information about participants and companies was replaced with tags such as [Company B], [Colleague A], and [Client C] to guarantee anonymity [27].

After the transcripts were finalized, the Dutch transcripts were translated to English using DeepL[11]. DeepL is a tool based on artificial intelligence in order to translate text. After using this tool, the author has reviewed and edited the content as needed and takes full responsibility for the content of the work. Whenever a quote is used, the English quote is

[11] https://www.deepl.com/en/translator

checked against the original Dutch quote and changes are made whenever the automatic translation is not entirely accurate.

Following transcription, open, axial, and selective coding is performed [22]. The stages in this process are outlined in Figure 14 and the process is performed solely by the author. To facilitate this process, the transcripts of the interviews were uploaded to Atlas.ti [105], which is a tool that can assist with coding. The process of open coding was initiated with the first transcript (INT1), from which an initial coding scheme was created. This scheme was expanded throughout the analysis of each transcript, up until and including INT8. For each of the interviews, the transcript was first fully read, while taking notes for codes on paper. In the next round of reading, paragraphs were highlighted, and codes were added in Atlas.ti. This is a form of open coding, where categories are created from concepts and variables [22], [25].

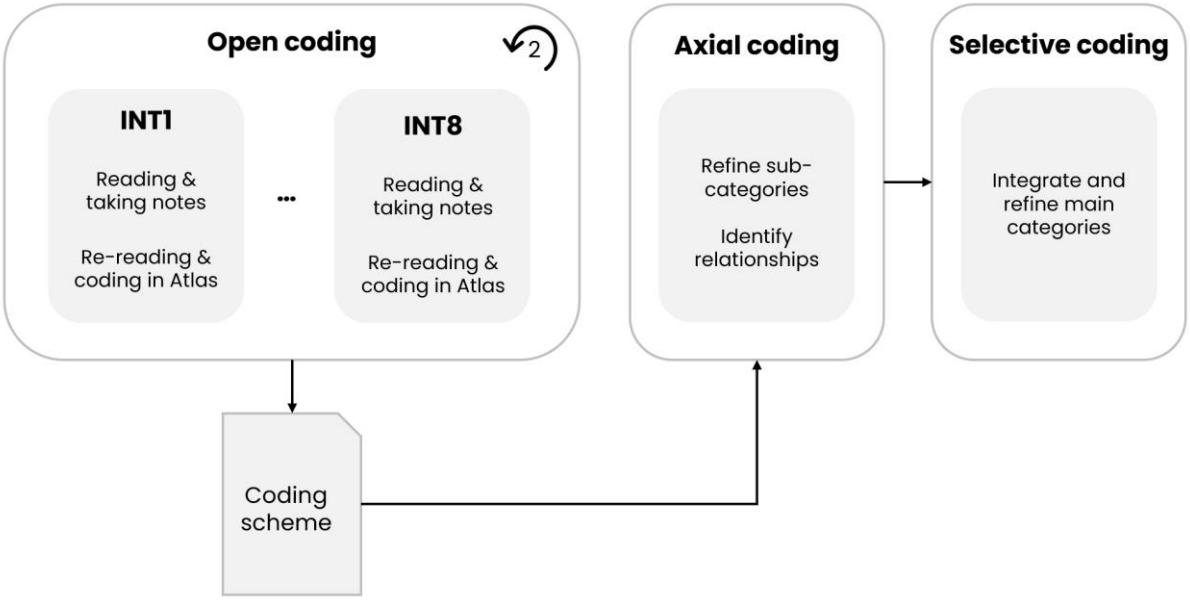


Figure 14 - Coding process

After all transcripts were coded once, a second round of coding was performed to check the first transcripts on codes that emerged later in the process. Subsequently, the sub-categories of the codes were refined through axial coding. Throughout this process relationships were identified as well. As a final step, selective coding was used to integrate and refine the main categories [22], [25]. The use of these distinct coding stages is necessary to ensure a rigorous and systematic analysis of the data, where each stage builds upon the previous one to gradually refine the data. The overview of identified codes and their description can be found in Appendix 9.7.

Following the coding process, themes were identified in accordance with the steps by Clarke and Brown [26], as described in Section 2.4. First, the codes were sorted into initial themes. Subsequently, these themes were reviewed and refined by evaluating whether the themes and data extracts form a coherent pattern. As a next step, the essence for each of the themes was defined, followed by the final step of producing the report. Figure 15 illustrates the identified themes (represented by white squares), their corresponding codes (depicted by coloured squares), and the relationships between the themes (indicated by grey lines). In accordance with phase 5, defining and naming themes, the description for each of the themes can be found in Appendix 9.8.

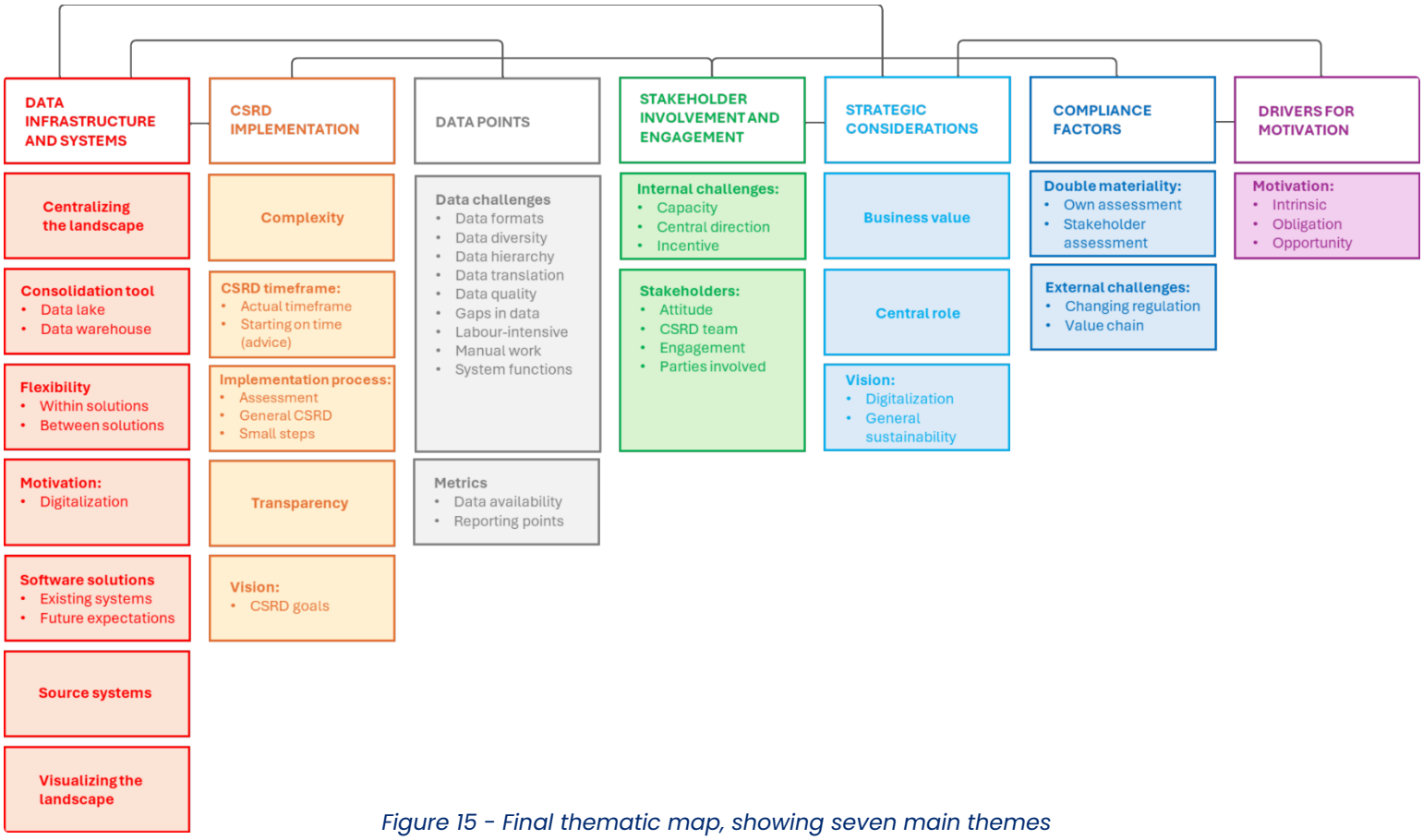


Figure 15 - Final thematic map, showing seven main themes

4.3 INTERVIEW RESULTS

Based on the answers in the interviews, this section presents the results of the thematic analysis based on the themes depicted in Figure 15. In Section 4.5, an overview is presented of the key findings that contribute to the development of the roadmap.

4.3.1 Drivers for motivation – Obligation, intrinsic, and opportunity

In order to gain insight into the attitudes of organizations towards CSRD, motivation is the first theme we discuss. It is prevalent among various organizations that CSRD is perceived as an obligation. This perception demonstrates itself in various ways in the responses of those interviewed, suggesting that compliance is often the driving factor in addressing CSRD requirements.

The resistance to engage beyond the necessary minimum is prevalent in some organizations ("*The bottom 5% [of employees in Company B] want to do nothing at all; only the pure necessary.*" [INT2]), while other groups of employees do see it as an opportunity ("*We're going to do more […] I think that it's also an opportunity. And also a commercial opportunity.*" [INT2]).

And while policy might not be inherently part of an organization, they do try to see the positive aspects of adhering to CSRD ("*We do try to see the positives in it […]. But [Company C] is not crazy about policy… CSRD runs a lot on policy, so that's […] a bit of a must.*" [INT3]).

For organizations with a long history of (voluntary) sustainability disclosures, such as Company E, CSRD is viewed as a "compliance exercise", because they make a lot of disclosures that are often more extensive or sophisticated than what regulations require [INT5]. And even though organizations experience the CSRD implementation as a must [INT3, INT7], some do consider it as highly important to implement the directive responsibly ("*Well, we have to do this, but we are going to do it right.*" [INT7]).

Considering the organizations that initiated the process out of intrinsic motivation, employees described it as a matter of voluntary engagement ("*You should do it because you want to, not because it's compulsory*". [INT4]). Furthermore, they found it surprising that the directive would not apply to them ("*We at [Company F] though that was very crazy […] and then we also said you know what, we'll just do it!*" [INT6]). Even the trade association representing the social housing sector in which Company F operates, has initiated the development of a standardized framework for all corporations within the sector. This framework can be used by organizations who wish to pursue the CSRD compliance.

Nevertheless, whether the driving factor is of obligatory or of intrinsic nature, organizations are motivated by their experience of viewing the CSRD as an opportunity. Some organizations express a desire to lead in their efforts ("*From a sustainability point of view, we always like to be at the forefront.*" [INT1]), while others hope CSRD will lead to change ("*To initiate a change and make people more aware.*" [INT4] and *"[…] they're looking to create more consistency, standardization in the way these disclosures are done and the quality of the data and so forth. So […] that part is to be welcomed*." [INT5]).

This is further elaborated on by interviewee 7, who sees the CSRD as an opportunity to adopt a new perspective and provide a clear direction for organizations to take in the upcoming years (*"[…] such a CSRD also forces you to think, well, it really does dust things off and you look at how you do things in a different way. […] I do believe that this is a basis that will enable us to accelerate together over the years. […] You need these insights in order to take action and not run around like a headless chicken."* [INT7]).

A final perspective highlights the opportunity that CSRD provides for organizations to showcase their successes and to focus not only on their shortcomings ("*[…] we can also show that we are doing certain things well. So, it's also an opportunity in that sense […]."* [INT6]).

4.3.2 CSRD implementation – Timeframe, implementation process, vision, complexity, and transparency

Moving on beyond motivation, the second theme discusses the process of working towards CSRD compliance including the timeframe, the vision for future steps, and companies' experience with complexity and transparency.

Timeframe

Starting with the timeframe for CSRD implementation, some companies have worked on general sustainability for a longer period, which then transformed into CSRD compliance within the last three years [INT1, INT2]. Interviewee 2 notes that despite early initiatives, competing priorities initially hindered substantial progress until external pressures urged further action about two years ago [INT2].

Company C began the process early last year with the help of a consultancy company ("*In early March was the kick-off with the consultant and then we also really started data collection, […] the double materiality matrix [and] stakeholder analysis.*" [INT3]). For Company E, CSRD was nothing new ("*[…] in general with a lot of these regulations, we've been following them for a long time.*" [INT5]) and the organization started focusing on CSRD early last year as well.

Making a swift start, Company G started their double materiality analysis already in 2022, facilitated by external consultants ("*Only at that time, of course, the requirements of the CSRD were not yet fully known and certainly not final.*" [INT7]). Therefore, they are now still finalizing the official analysis. Interviewee 8 describes a relatively recent entry into CSRD activities, starting at the beginning of 2024. Similar to Company C, they sought help from a party to provide them with some structure ("*Where do you start and where do you end?*" [INT8]).

Considering these timeframes, each organization that is part of a mandatory reporting group expressed the importance of starting on time. Interviewee 1 explained that they gradually learn what data to record and collect, while it takes a long time until everyone is "*looking in the same direction*" [INT1]. Organizations should not underestimate how much work it is to comply with the CSRD [INT3], which was also expressed by Interviewee 8 ("*When we were just starting it, we didn't have a clear vision of how big it would be, how extensive. So please start on time.*" [INT8]). The same issue arose for Company G, as they spent quite a long time validating the double materiality assessment (DMA). This in turn came at the expense of time to really implement things [INT7]. Furthermore, it is important to have a clear scope and build internal capacity, as with any project [INT5]. Interviewee 2 found an all-encompassing quote for the implementation process: "'*CSRD is like a marathon, start training on time'. […] You don't just do it on the side.*" [INT2].

Implementation process and vision

Focusing on the CSRD implementation within these timeframes, organizations have followed various methodologies and processes. Company B focused on first creating support within the organization, while simultaneously performing baseline measurements (*"[…] we did the Ecovadis assessment, […] product LCA*[12] *[…], and company LCA…"* [INT2]). They are currently in the process of the DMA, aiming to be capable of reporting by the end of 2024. Based on their experience, Interviewee 2 advices to have a clear implementation plan as they are concerned Company B will not meet their deadline to start reporting from January 2025 on (*"[…] create a clear roadmap that you can stick to, because before you know it, you're back to […] working on the day-to-day business."* [INT2]).

Company C is at a similar stage for the DMA, after first performing a stakeholder analysis. An additional step was to perform a gap analysis regarding their data including advice for improvement (*"[…] that shows where you can find what data and how reliable it is or if there is still something missing."* [INT3]). For them, the focus lies on finalizing stakeholder conversations as well as defining what systems to get ready.

Interviewee 4 discusses the struggle to cluster material themes and present them effectively (*"We actually don't really have a process […] because we have no obligation."* [INT4]). They completed the value creation model and stakeholder analysis but have not yet been able to identify how they will report on the material themes. However, their aim is to be well prepared once the assignment comes or they need information to be ready for other stakeholders.

Focusing more on the technical aspects, Company E started building the infrastructure middle of last year. In 2024, the focus lies on starting data collection in areas where this was previously lacking and ensuring this data is of high quality (*"We are of course subject to assurance now."* [INT5]). From next year on, they will start thinking about the longer term in terms of IT, once they have finalized the DMA (*"We need to […] start defining the business requirements."* [INT5]). Interviewee 6 explained that they are in a similar technical phase, checking how much data is available for each data point. They have not yet finalized an official DMA, but that is scheduled for the second half of 2024 [INT6].

Company G is working on finalizing the official DMA, while at the same time focusing on the more technical part as well as data collection (*"To ensure that you do that in the same way, we have created a 'Definition of Done' through which we take all those points."* [INT7]). A technical gap assessment based on the 2022 annual report has shown them where metrics or text is missing to be CSRD compliant, so that is where the focus currently lies. The aim is to have all data points ready in September, such that a first version of the annual report can be created. Afterwards, the focus should shift to the phase-in requirements that will be introduced after 2024 (*"We have completely ignored those for now."* [INT7]).

[12] Life Cycle Assessment

Company H has had a slightly different approach thus far, compared to the previously discussed companies. After selecting a partner, the second step was to start mapping out the CO_2 footprint for their whole group. That should be finished in July 2024, after which they will start the formal process of the materiality analysis. The aim is that they can define in autumn what data they want to start measuring and what the necessary resources will be to commence reporting (*"[…] early next year we should be able to report fully on the first dataset based on that."* [INT8]).

Based on their experiences, five of the interviewees recommended to do the implementation in small steps (*"Take one step at a time and don't try to do it all perfectly at once, because you won't make it anyway."* [INT2]; *"Celebrate small steps, be happy also with small results."* [INT4]). Just making a start, breaking it down, and tackling everything piece by piece is believed to be the most effective way to undergo such a process [INT3, INT6, INT7].

Complexity

When discussing the complexities associated with implementing the CSRD, interviewees frequently described the process as *"vague"* [INT3] and *"very complex"* [INT1, INT6]. Interviewee 1 added the challenge of *"keeping it fun"*, noting the difficulty in keeping the process enjoyable [INT1]. Interviewee 2 explained that the complexity lies in the novelty of the topic (*"Well, quite complex. Because it is a new subject matter, and all the more so because it has not been normalised enough yet, so it can be interpreted and implemented in different ways."* [INT2]).

Interviewee 5 expressed concerns regarding the understanding of the directive, particularly for smaller companies. They remarked (*"[…] if a large company like [Company E] with 25 years or more of sustainability reporting experience can't understand and implement the CSRD, then […] the middle-sized companies in [e.g.] Germany have no hope of being able to do that."* [INT5]). They did add that you *"can't let the perfect be the enemy of the good"*, because there is a lot which is unclear. Therefore, companies should aim for *"good enough"* when it comes to the first reports [INT5].

Considering stakeholders, the representative for Company G emphasized the extensive scope of the CSRD, covering multiple departments as well as a stretch towards customers (*"[…] it affects your whole company and environment, social, and governance."* [INT7]). Furthermore, the complexity extends to stakeholder engagement, with specific reference to the questions that need to be asked to stakeholders such as employees or suppliers. Interviewee 8 questioned the clarity of these inquiries and whether the answer will even be usable (*"[…] there are questions in there… I first have to read it three times myself and what does it actually say? What do they actually mean? And that's what you then require your employees, your suppliers to fill in."* [INT8].

Transparency

Alongside complexity, several interviewees articulated the importance of transparency in reporting, particularly in relation to the data presented. Company A highlighted the need for honesty regarding anticipated reductions in CO_2 emissions over the years; they acknowledged the importance of not overstating achievements ("*Yes, we are going to save 50% [on emissions], but we are not [actually] going to save 50%; we are going to estimate 50% more accurately. And I think we have to be very careful that we don't claim '[…], we are emitting 50% less CO_2' while in reality nothing has changed.*" [INT1]). Furthermore, Interviewee 1 noted the desire to ensure that actual sustainable changes are clearly visible for their stakeholders.

Company B initiated the creation of Environmental Product Declarations (EPDs), which serve as a "passport" for their products, detailing aspects such as materials used, the percentage of recyclable material, and CO_2 emissions [INT2]. Interviewee 4 pointed out that CSRD-reporting should not become some kind of "*idealistic story*", stressing that companies must be honest about the motivator behind their decisions [INT4].

Interviewee 7 added that this level of transparency will place new demands on employees, for which some resistance was observed ("*'But I just want to do nice things.' No, now you have to write it down, show it and be transparent about what you do. That won't always be perceived as fun by everyone.*" [INT7]). However, they also expressed optimism that the ability to compare data between companies can act as an accelerator for collective process in ESG issues. They noted that for the first time, ESG issues are being addressed with data through quantitative metrics, which they believe will lead to greater clarity.

4.3.3 Compliance factors – Double materiality and external challenges

As already touched upon in Section 4.3.2, the implementation of the CSRD involves several compliance factors, notably the double materiality assessment. When discussing this step of the process, it became apparent that multiple organizations first conduct their own (internal) version of the DMA, given the lack of clear guidance during the early stages of the CSRD's rollout. Company B, for example, began prioritizing relevant themes in 2023, anticipating the official DMA guidelines [INT2]. Based on this assessment, the focal points are currently known, and they can start with this direction as the expectation is that the outcome of the official DMA will not be very different ("*[…] it will come out 90% the same.*" [INT2]).

Similarly, Company C created an initial materiality matrix by rating themes from a longlist provided by the NBA[13], which facilitated the development of a pilot CSRD report. Their experience indicates a significant overlap between their internal assessment and the expected official materiality matrix [INT3]. In another instance, Company F initially chose to engage with the 17 Sustainable Development Goals to determine material themes, finding it "*too much*" to discuss the entire longlist with stakeholders at that time. This year, they plan

[13] Koninklijke Nederlandse Beroepsorganisatie van Accountants

to undertake a comprehensive DMA [INT6]. Company G commenced their DMA process in 2022 and is currently striving to align with CSRD requirements [INT8]. Company H, having also conducted an internal assessment, neither anticipates significant deviations in the official DMA (*"[…] we ourselves I think know best what is or is not material."* [INT8]).

Considering the compliance factors in general, the interviewees highlighted several challenges. First and foremost, the challenge of changing regulations, which adds to the complexity of the implementation process. Interviewee 2 emphasized that even parties specialized in CSRD implementation provide different advice in 2024 compared to 2023 when it comes to the same question (*"Everyone is learning; things are changing enormously."* [INT2]). The compliance process therefore requires a pragmatic approach [INT2]. Company H reported similar issues, noting that their advisory partner often encounters unforeseen questions or issues (*"[…] you notice that it's all new for them too."* [INT8]).

For Company E, the dynamic regulatory environment presents additional challenges, particularly due to the difficulty in retrofitting legacy systems, which is both time-consuming and costly. Interviewee 5 criticized the unstable regulation and inconsistent guidance from the EU, which complicates the company's ability to commit to specific systems (*"It becomes actually very difficult for us to make commitments to systems because the EU doesn't follow a predictable process."* [INT5]). This learning effect is also apparent for Company G, who highlighted that the learning curve of this novel topic not only affects companies working on CSRD compliance, but also auditors and EFRAG itself [INT7].

Furthermore, companies face external challenges related to the upcoming requirement for data collection across the entire value chain. Interviewee 6 identified this as a primary challenge for most companies, especially concerning the estimation of end-use data [INT6]. They additionally reported difficulties for Company F associated with changing stakeholders, which complicates both the data collection process and the consistency of the data gathered (*"[…] it's not the case that you always perfectly work with the same parties, which is kind of assumed with something like this [CSRD], so that you always get the data in the same way."* [INT6]). Additionally, Interviewee 8 pointed out the challenges of dealing with "*different systems and different languages, different people…",* describing the process as a significant "*hassle*" so far [INT8].

4.3.4 Strategic considerations – Vision, central position, and business value

Moving forward, the CSRD presents not only challenges but also strategic opportunities for organizations. These are reflected in overarching visions concerning sustainability and digitalization; having a central strategic position in a sustainability context; and sustainability being a driver to create business value.

Vision

For numerous companies, sustainability is integral to their corporate strategy and vision. Company A has long prioritised sustainability, consistently trying to position themselves at the forefront of this issue [INT1]. Similarly, Company B has made sustainability a priority in their operations, developing an ambition, vision, and mission, supported by ESG targets (*"[…] our mission is to go and fill that pioneering role."* [INT2]). The data collection mandated by the CSRD serves as a baseline measurement in this regard. Company G also considers themselves at a *"front runner position",* having engaged with environmental issues for many years [INT7]. According to Interviewee 4, sustainability is very close to the heart of Company D, having incorporated the ESG topics in the core of their mission, vision, and strategy [INT4]. Likewise, Interviewee 8 expressed pride in Company H's long-standing commitment to sustainability over the past 10 to 20 years [INT8].

Despite not being obligated to comply with CSRD reporting requirements, Interviewee 6 remarked that it would be *"heartily crazy"* for Company F not to report on its sustainability initiatives, given that sustainable initiatives are at the core of their work as housing association [INT6]. However, Company C finds it challenging to align the CSRD with their existing sustainable practices, which they have been working on for years (*"[…] it is a very sustainable building, but that has nothing to do with the CSRD. […] We are quite sustainable, but we do it because it is the logical thing to do, not because the CSRD requires it. There's a bit of resistance there."* [INT3]).

Considering digitalization, organizations share a common objective of advancing their capabilities, albeit with varying strategies. Company A aims to connect all suppliers to a visibility platform, establishing a single source of truth for mapping sustainability and visualizing transport movements [INT1]. Company B similarly aspires to maintain a single truth for all stakeholders (*"[…] uniformity is really a must."* [INT2]), although they do not wish to organize this themselves. Interviewee 5 acknowledges the need for system changes, with a wish for more consistency across systems, though Company E is still evaluating over the next four years whether to implement a data lake, data warehouse, or another solution [INT5]. Company F envisions the use of a comprehensive system, similar to those used for financial records (*"That's the best and that's the ideal picture."* [INT6]). What that exactly looks like is still unclear. Company G seeks to digitize processes to minimize errors, with plans to explore software tools in the coming years, as Interviewee 7 acknowledges the need for change [INT7]. Finally, Company H intends to utilize the sustainability management platform developed by Coolset as much as possible, supported by a data warehouse and Power BI reporting. However, their immediate focus is on capturing only relevant data [INT8]).

Central positioning

Several companies have identified a strategic opportunity in assuming a central position in the CSRD process, acting as a key partner for their stakeholders. Interviewee 1 explained that Company A serves as a "*collection point or platform*" for all international suppliers, handling compliance for CSRD as well as other legal requirements (*"We'll take care of it, just like we already do with excise and import taxes."* [INT1]). Initially, Company A looked to their stakeholders to evaluate their actions, only to find that those stakeholders were awaiting Company A's initiative (*"[…] we are definitely in a director's role there, and I hadn't seen it that way at first."* [INT1]). Company D similarly aims to adopt a proactive position *"as a big battleship*", rather than being reactive [INT4]. Likewise, Company G expects many companies to *"knock on their door"*, having a central role as a logistics service provider [INT7].

Business value

Apart from having a central position in their supply chain for data collection, the business value of sustainability emerged as a key theme in discussions about the CSRD. Company A identified that insights from a sustainable viewpoint can lead to actions that result in cost savings. For instance, reducing the weight of the glass of wine bottles enables more efficient shipping, thereby reducing both costs and emissions (*"If you can put an extra pallet in a shipping container, you can easily achieve an extra 5% in both costs and emissions."* [INT1]). This is also valid for their transport via different means. Furthermore, the creation of a central sustainability platform offers a unique selling point, facilitating stakeholders' reporting processes.

Interviewee 2 noted that for Company B, increased insight into the materials used in their products can lead to innovations in product design and reuse. This presents a significant commercial opportunity in form of potential profit; especially given the volume they transport [INT2]. Interviewee 3 suggested that effective CSRD reporting can be an opportunity for companies that are currently not doing well in terms of profit (*"[…] then you can hopefully keep your image afloat […]"* [INT3]).

Interviewee 6 highlighted an interesting opportunity for Company F, explaining that sustainability KPIs can result in discounts on bank loans (*"[…] we set up KPIs and we link them to our loans, and the moment we achieve those KPIs we get a discount."* [INT6]). Company H enjoys similar benefits, with interest reductions on loans for new sustainable office premises. Additionally, Interviewee 8 pointed out that questions about emissions and targets are increasingly common in tenders, with the expectation that this will play *"a crucial role in winning tender contracts"* [INT8]. Therefore, it is an additional reason for them to work on their CSRD reporting to be able to provide answers to such questions and enhance their competitive edge.

4.3.5 Stakeholder involvement and engagement – Stakeholders and internal challenges

One of the main aspects of CSRD implementation is the engagement with stakeholders, both on an internal and external level. Discussions need to be held with stakeholders in the value chain, but additionally external parties may be involved in advisory roles. All these stakeholders have their own attitude towards CSRD implementation. Additionally, internal challenges can arise related to capacity, a lack of central direction, and the challenge of creating incentive.

Stakeholders – CSRD team

The team structures responsible for implementing the CSRD vary significantly across the interviewed companies. Company A utilizes a dual-layer approach where the global and national sustainability teams are in charge, while smaller teams such as inbound logistics are responsible for their own data gathering [INT1]. Company B initially handled the CSRD through the innovation manager, who managed the process on the side. However, a dedicated CSRD lead has been appointed and multiple departments and employees are involved in the process such as purchasing, logistics, product development and management, HR, and the CEO [INT2].

Company C has established a dedicated CSRD working group, including five employees, two graduate students, and an intern. The group's composition includes a representative from marketing, finance, HR, procurement, and the company director [INT3]. Similarly, Company E, being a larger organization, employs a larger working group with about 15 to 20 core members and an extended team of 100-150 people across the organization. This group includes eight experts in double materiality, an ESG data controller, a lead on sustainability-finance interaction, communication, controls, and IT. Each CSRD topic additionally has a designated owner, emphasizing the importance of leveraging internal expertise (*"[…] in a lot of areas it's actually […] better for us to take control of it ourselves and figure it out, rather than [hiring] a consultant."* [INT5]).

In Company D, the CSRD team consists of individuals specializing in environmental issues, Human Resources (HR) for social aspects, and governance experts. A core team, including one or two employees from each service, collaborates on the annual report, supported by additional staff focused on detailed data collection and cooperation with the internal BI unit to address technical challenges. This collaborative approach helps overcome complexity, as expressed by Interviewee 4 (*"[…] do so as a joint effort, in which all voices are heard and can be included."* [INT4]).

Company F initially placed the CSRD process within the finance department, involving three business controllers, the manager of finance, and a financial controller. Additional colleagues, such as a manager for strategy and innovation and a sustainability manager, are involved per section of the CSRD, emphasizing the importance of broad internal stakeholder engagement. They suggest to *"involve every department, so it shouldn't become a party of finance or a party of real estate. It is something that affects a lot of areas in the organization"* [INT6].

Company G organizes their efforts around a three-pronged structure: the ESG department, reporting/internal audit and risk management, and the business unit. Interviewee 7 considers the inclusion of individuals with legal and regulatory expertise to be of great value [INT7]. Finally, Company H currently has a smaller CSRD team, comprising the head of finance, the head of HR, and the office manager, with plans to involve a working student to enhance the process focus from mid-July [INT8].

Internal challenges

Considering the internal stakeholders, several challenges were identified by the interviewees. The first we discuss is capacity. Interviewee 2 noted that their organization has been engaged in the implementation of a new ERP system for the past year, which has taken longer than anticipated. This resulted in many stakeholders dedicating up to four days a week to this project, thereby limiting the capacity available for CSRD initiatives (*"So the plan is there, only we can't really make metres now, so we're a bit behind…"* [INT2]). Similarly, Company D faces prioritization issues, struggling to free up resources amidst competing priorities (*"Right now we have a slightly reactive attitude there, because it's just with other priorities…"* [INT4]).

Likewise, at Company H, CSRD responsibilities are not the primary focus of the CSRD team members, leading to its lower prioritization (*"That is also a bit of the pitfall […], because it is often a bit lower on the priority list."* [INT8]). Despite "*good intentions*", there is a lack of structural focus, with employees unable to dedicate one or two days a week consistently to CSRD activities. For Company G, the capacity issue is mainly related to IT, as Interviewee 7 expressed "*you always have to battle against all the other projects out there"* in terms of securing budget and IT specialists [INT7]. Interviewee 6 suggested that larger companies might need to hire additional staff to address these challenges [INT6].

Another significant challenge identified is a lack of a central direction to follow within the organization. Interviewee 4 emphasized that while numerous CSRD-related activities are occurring simultaneously, there is an absence of central direction (*"[…] the integral direction is missing."* [INT4]). Interviewee 5 highlighted the difficulty in comprehending both the regulatory framework and the IT landscape, which complicates the alignment of internal processes. They suggested that clearly defining the business requirements could help "*generate internal alignment"* and achieve a common understanding [INT5].

Furthermore, creating incentive to participate in reporting tasks and making the compliance process engaging presents a significant challenge. Interviewee 1 pointed out the difficulty in motivating employees, noting that a creative approach is required to make CSRD compliance "*fun"* beyond the necessity of the task (*"How do you ensure that there is also a good incentive to get going with it, other than that 'it has to be done'?"* [INT1]). They believe it is important to make employees aware that CSRD compliance can contribute to better plans, save money, and introduce interesting initiatives. Similarly, Company B identified "*getting people on board"* as their main challenge [INT2]. For Company C, the issue revolves around the perception of reporting; employees prefer focusing on doing and improving, rather than documenting processes (*"That something has to be done when it*

doesn't actually add any value." [INT3]). Company H is in a likewise situation; while there is enthusiasm for sustainability initiatives, the compliance aspect is often seen as burdensome (*"[…] it's a horror in that sense."* [INT8]).

Stakeholders – Parties involved, attitudes, and engagement

The involvement of stakeholders and external advisory parties plays a crucial role in the CSRD implementation process, albeit those often have different attitudes towards the new directive. For the inbound logistics department of Company A, the main stakeholders are their hauliers. These companies have various levels of motivation regarding CSRD reporting, but Interviewee 1 has identified a central need for a standard or central platform. They particularly noticed transport companies tend to have difficulties with the transparency required by the CSRD (*"Because that is partly the blacksmith's secret. That's their business."* [INT1]).

Company B collaborates with Use Impulse, a company that takes care of the whole CSRD process, and Novel-T[14] for innovation, emphasizing the importance of collaboration (*"You can't do it alone, you really have to work together. Be open to working together."* [INT2]). They engage with key stakeholders, including three significant customers and two major suppliers (*"They [the customers] are actually ahead of the regulations. Those have also built their own ESG structure."* [INT2]). Company C has a similar partner; a consultant from Waardevol MKB for general CSRD guidance. Their stakeholder engagement includes a range of customers and suppliers, each with varying levels of data availability. Interviewee 3 noted the importance of early engagement with an accountant, as delaying this aspect led to unexpected disclosure requirements (*"So now I have to start thinking about what we did last March."* [INT3]). Regarding their value chain, Company C seems to be ahead of their partners, with some even charging *"€3,000"* to collect data for them.

Company E adopts a more self-reliant approach, utilizing consultants selectively but generally preferring to handle tasks internally. However, they recognize the value of external perspectives gained at other companies, particularly for IT-related issues [INT5]. Similarly, Company F manages the CSRD process internally, while engaging with stakeholders such as tenants, municipalities, cooperation partners, suppliers, banks, and supervisors [INT6].

Company G seeks assistance from PwC, especially for interpretation issues, while also consulting with stakeholders including employment agencies and delivery partners. They exchange data on metrics such as driven kilometres, vehicle types, and emissions [INT7]. Lastly, Company H has partnered with Coolset, a company providing a CSRD data tool and implementation guide based on questionnaires. Interviewee 8 underscores the value of such partnerships to provide structure (*"[…] if you start inventing the wheel all by yourself, you won't go anywhere."* [INT8]). Their value chain stakeholders include three major customers, three main suppliers, and employees. Interviewee 8 notes that within their

[14] Novelt-T is a non-profit organization that challenges entrepreneurs with innovative ideas to start, innovate, and grow.

group, which comprises companies outside of Europe, there is often less emphasis placed on the CSRD process, creating challenges in achieving process ("*They say: 'what are you worried about'?"* [INT8]).

When engaging with stakeholders, creating a sense of urgency is considered crucial by Interviewee 2, emphasizing the discussion of sustainability with all internal stakeholders ("*Creating support is super important; explain, explain, explain."* [INT2]). They believe that taking a pioneering role involves engaging with customers and partners, giving an example on their cooperation regarding sustainability issues with a competitor. Similarly, Company H initiated a general internal kick-off to underscore the importance and needs for the CSRD process [INT8].

Company C is in the early stages of collaboration, with initial questions being exchanged with one of their stakeholders [INT3]. Company F has progressed further, asking stakeholders what they require from the organization, while also needing information from demolition partners about the reuse of materials. Additionally, one of their financial lenders has increasingly asked about their practices. Internally, Interviewee 6 highlights the necessity of making CSRD processes accessible and manageable for employees ("*So that you sort of translate it towards the department."* [INT6]).

4.3.6 Data infrastructure and systems – The landscape, source systems, software solutions, flexibility, and consolidation tools

The implementation of the CSRD requires a proper digital infrastructure. This section examines companies' motivations for digitalization within the context of CSRD, the various source systems utilized, the visualization of their IT landscapes, software solutions for CSRD reporting, flexibility of systems, and the use of consolidation tools.

Motivation – Digitalization

Companies generally exhibit a strong motivation toward digitalization, recognizing its importance for the visibility of sustainability. Interviewee 1 noted that Company A aims for a centralized digital landscape using a *"limited number of strategic carriers"* and organizing a digital platform to serve as a Single Source of Truth (SSoT) [INT1]. For Company B, software and digital tooling are not in their own scope, but these matters are outsourced to partners. The organization desires an SSoT but does not consider this to be their own responsibility [INT2]. Interviewee 4 emphasized to *"go above and beyond on a technical level"*, as ultimately *"what you put in is what you can get out"* [INT4]. Companies F and H are considering central systems but have yet to take specific steps to go there [INT6, INT8]. Company G seeks to digitize as many processes as possible, from figures to action plans [INT7].

Source systems

To compile data for CSRD reports, companies utilize various source systems. Many work with Enterprise Resource Planning (ERP) systems in which a lot of information is available [INT1, INT2]. Examples of ERP systems used by the companies in this study are AFAS [INT6] or SAP [INT7]. Others rely more on manual data gathering from sources like supplier invoices, annual statements, and HR systems [INT3]. Company D, situated in the education sector, has a primary system for student affairs but supports it with many additional applications [INT4]. In the case of Company E, most data comes from five or six major applications, focused on topics such as safety, environmental issues, finance, and HR [INT5]. While ERP systems can be used as a central base, often other systems are used simultaneously; companies F and H connect their ERP systems to Power BI and additionally use Excels for calculations [INT6, INT8].

Visualizing the landscape

Visualizations help organizations map their IT landscapes, providing an overview of source systems and data. Company A, for example, worked with university students to develop enterprise architecture models, gaining insights into the applications of their inbound logistics department [INT1]. Having made more of a list-overview, Company C identified where what data can be found and whether the source is reliable [INT3]. Company D mapped underutilized data sources and where they can improve compliance [INT4]. Company E uses CSRD data points as a foundation and tracks various aspects, such as definitions, boundaries and underlying source systems [INT5]. Company F has an enterprise architecture model and additionally created models of the data process flow for each data point (*"What is asked here, this comes from there, and this system touches that, […]"* [INT7]). Although Company H has yet to create an official overview, they recognize the value of doing so (*"That's actually a good tip."* [INT6]).

Software solutions

A range of software solutions is utilized for CSRD reporting, including existing systems adapted for new purposes and tools specifically designed for CSRD. However, the expectations are that many developments are still to come. Company B began with Excel to start small, in combination with an LCA programme to conduct the assessments [INT2]. However, in Interview 3 it was highlighted that Company C's auditor classified Excel as *"not reliable"*. Company C still uses Excel sheets for now but acknowledges its limitations. As their core application, they currently use Milieubarometer for environmental data [INT3].

Company A uses Big Mile for CO_2 calculations, Shippeo for transport visibility, and they will invest in a product information management tool (*"To be able to record more clearly at article level as well, what are all the different characteristics of an article."* [INT1]). The carbon platform by Realised is used by Company D for Scope 1, 2, & 3 calculations [INT4]. Company E employs Workiva, a cloud-based collaboration tool that can link data tables within reports, alongside various spreadsheets due to data hierarchy complexities [INT5].

Considering that many CSRD systems *"are mainly a place to store data and then get a graph"*, Interviewee 6 uses a Microsoft Teams environment, somewhat functioning as a

data warehouse, to upload relevant data (*"I would like to have a system, but then it would have to add value, […]"* [INT6]). Company G uses SAP as a basis for ESG data, but still relies on Excel for some tasks, while Nossa Data facilitates collaborative report writing [INT7]. Company H utilizes the software platform of Coolset to track CO_2 emissions and guide their CSRD implementation, although this currently requires manual data input [INT8].

The future of these systems is expected to evolve significantly as the field matures [INT2, INT6, INT7]. Interviewee 2 anticipates a universal smart database to streamline ESG reporting (*"From the European Union, of course, there is a desire to make ESG universal."* [INT2]). Interviewee 6 expects systems like AFAS to integrate reporting features over time, expressing that *"things like that will come as they go along"* [INT6]. Company G plans to wait for further developments before fully committing to new tools, seeking to avoid the initial *"pains"* of current tools and systems. Nevertheless, especially data exchange should be done through a system to provide a quality guarantee (*"You don't want that [data exchange] via an email or whatever."* [INT7]).

Flexibility

Flexibility is a consideration for organizations when investigating these software solutions. Interviewee 1 highlighted the importance of being able to adapt to new technologies and market entrants; so, starting early to gain experience with the possibility to stop and switch (*"[…] deploy the solutions in such a way that they can also move along […]"* [INT1]). Interviewee 2 emphasized the need for pragmatism in navigating the rapidly changing landscape [INT2]. Additionally, systems should be flexible enough to accommodate the diverse data fields required by the CSRD and other regulations, as these requirements tend to emerge *"rather shock-wise"* [INT1].

Consolidation tools

As a basis for the software solutions, several companies explore the use of data warehouses or data lakes as foundational elements for their reporting systems. Interviewee 1 linked a data warehouse to flexibility and efficient reporting, highlighting their ongoing efforts [INT1]. Company D utilizes both a data warehouse and a data lake for their dashboards [INT4]. Likewise, Company G stores several types of data about employees, premises, vehicles, and customers in data lakes [INT7]. Company E is considering a data warehouse or lake after addressing other challenges (*"In a way right now I'm probably more focused on hierarchies."* [INT5]). Similarly, Company H is preparing to integrate their group's ERP systems within a data warehouse, aiming for comprehensive reporting (*"[…] we can generate our reports for the entire group via Power BI […]"* [INT8]).

4.3.7 Data points – Metrics and data challenges

Based on the infrastructure and source systems, data points are collected for the CSRD reporting. However, several challenges arise regarding this data.

Metrics

Considering the data points, each organization selects and prioritizes these based on their materiality assessment. Interviewee 1 emphasized the importance of emissions for their department, which are calculated by considering factors such as product movements, weight, transport type, and container type [INT1]. Company B views ESG to go beyond carbon emissions, highlighting the significance of sustainable entrepreneurship. Having a physical product, their Scope 1 and 2 is within their own production capacity, but the end-use is a significant topic for their products and therefore included in S4 – consumers and end-users (*"We actually do that [considering the end-user] very well already, but don't ventilate it yet, so I think that's another opportunity."* [INT2]). Company C primarily focuses on the environmental pillar, with additional data on the social pillar, considering metrics like consumption, material use, kilometres driven, waste, turnover and employee metrics. However, no specific KPIs and targets are set yet [INT3].

For larger organizations, almost all themes tend to be relevant (*"By the time you achieve a certain scale as a company, it is very difficult to say that data points are not material."* [INT5]). For Company E in the oil and gas industry, climate issues are of critical and strategic importance, while other aspects are managed on a day-to-day basis [INT5]. Company D also reports on most themes within each pillar [INT4]. Interviewee 7 noted that they have been reporting on climate change and emissions for years, but the CSRD requires slightly different summations or units [INT7]. Company G has four major company-wide KPIs related to ESG themes but lacks KPIs for all material themes (*"Circular Economy, […]. We don't really have a KPI on that one yet, also because it is a relatively new topic."* [INT7]). Their material themes include E1 – climate change, E2 – pollution, E5 – resource use and circular economy, S1 – own workforce, S2 – workers in the value chain, S4 – consumers and end-users, and lastly G1 – business conduct [INT7].

Company F, not being required to comply with the CSRD, has decided to begin small with three themes deemed material: E1, S1, and G1. They plan to expand their reporting once their official DMA is finalized [INT6].

The availability of these metrics varies significantly between organizations. Interviewee 1 acknowledged the presence of data but highlighted the challenge of identifying incorrect data points [INT1]. Company B estimates their data availability at about fifty percent [INT2]. Companies C and E both think that a lot of data is there [INT3, INT5], though Interviewee 5 noted that it is difficult to be certain (*"[…] the regulation as written is very difficult to implement."* [INT5]). For Company D, one of their lessons so far is that much of their data is already available, albeit scattered across various sources [INT4]. Interviewee 7 stated that Company G has all requested data available, but the CSRD requires a different method of consolidation and documentation compared to other reporting [INT7].

Company F has most of their data, except for some smaller data points such as employee training hours. Nonetheless, Interviewee 6 anticipates challenges in obtaining data from stakeholders in their value chain [INT6]. Interviewee 8 admitted uncertainty about data availability, since they have not yet completed their DMA (*"Well, to be quite honest, I have no idea. I think so, […]"* [INT8]). They consider that most of their data is financially focused, with a probable deficiency in operational data [INT8].

Data challenges

While collecting data for these metrics, several challenges arise for organizations. First, the diversity in data sharing formats and levels presents significant challenges. For instance, stakeholders vary in their methods of data calculation, with some averaging monthly, others per trip, and others across all transport (*"[…] everyone does it in their own way and that's the biggest problem for us."* [INT1]). Similarly, Company C experiences that for example plastic is measured by piece in some cases and by kilogram in others [INT3]. Company D, involved in education, faces the challenge of the difference between the standard calendar year and the academic year [INT4]. Interviewee 7 highlighted internal issues where different business units employ different systems, emphasizing the challenge of achieving uniform data for accurate summation, while most data is available (*"It's much more […] getting uniformity in that, so that you add apples and apples and apples together. That's much more the challenge."* [INT7]). For Company H, the fragmentation across the group further complicates data management (*"It would have been much easier if you had one company, with one location, with one system…"* [INT8]).

Specific software solutions require particular data formats, often needing manual adjustments. Interviewee 1 mentioned that data must be submitted into Big Mile in exactly the correct format before the answer rolls out (*"[…] Big Mile is, with all due respect, more just a really big Excel sheet."* [INT2]). For Company H the work is mainly on rearranging columns and fields, but that has been manageable so far [INT8]. Interviewee 6 noted that while employee data extraction from the HR system is straightforward, conversion into the desired reporting format is necessary [INT6].

The main challenge for Company E is that their systems are designed for operational purposes to manage workflows, where reporting is just an additional use case built on top (*"They weren't designed […] for the type of reporting that we're now having to do with them. They use different data hierarchies."* [INT5]). The misalignment of data hierarchies complicates exporting data into a data lake for manipulation, making that an impossible use case for now (*"[…] if our hierarchies all were aligned, it would make a lot of things easier."* [INT5]). This challenge was unique to Company E.

The translation of information into usable data is challenging for both Company B and Company F. Interviewee 2 explained that their ERP system contains extensive information, being the backbone of the company. However, improper data entry makes data retrieval difficult to perform *"flawlessly"* [INT2]. Interviewee 6 admitted *"a tonne of CO_2 tells me nothing",* struggling to comprehend whether this is much or not and what a certain amount of emissions looks like (*"[…] but we all still have to get that feeling […]"* [INT6]).

Data quality is another significant challenge. Interviewee 1 affirmed that ensuring data quality is not "*rocket science*", but maintaining high standards is difficult. Company A lacks data from before 2019, therefore requiring estimates that affect the accuracy of future emissions reporting ("*[…] higher quality of data automatically also results in fewer emissions.*" [INT1]). Moreover, varied levels of stakeholder data complicate error detection, making future-focused data quality efforts essential ("*[…] but what should I record?*" [INT1]). Interviewee 1 admits they do not always do this enough. Company B anticipates that initially, data quality will be poor, recognizing that the "*first year [of reporting] is not 100% for anyone*" [INT2].

Incomplete data also poses issues. Company C's assumptions, such as those based on petrol costs, highlight these gaps ("*[…] business traffic is completely based on an assumption around the cost of petrol. We don't have mileage […]*" [INT3]). For other values, Interviewee 3 must verify the numbers against invoices, a process in which overlooking an invoice is an easy mistake. Interviewee 4 highlighted the complexities in defining data quality standards to ensure objective reporting ("*[…] without giving a certain colour of flavour to what you present.*" [INT4]). Company E focuses on newly collected metrics, with each of them having an owner and specific definitions for scope, boundaries, and collection and calculation methods to maintain quality [INT5].

Data gaps further hinder reporting. Company C is unable to meet current reporting requirements due to missing data but aims to improve by 2025 [INT3]. Similarly, Company F acknowledges incomplete data but made the first step and is now working towards improvements [INT6].

Finally, manual data handling is error-prone and labour-intensive. Interviewee 3 described manually checking numbers regarding paper and raw materials on invoices, risking errors ("*[…] it could very well be that I overlook an invoice.*" [INT3]). Company D faces similar issues with manual data entry, expressing that "*there are in as many as 1,000 ways that can go wrong*" [INT4]. Company H also deals with manually copying meter readings into the CSRD tool to assess energy consumption [INT8]. Although Company E avoids manual data entry, they manually consolidate reports from various systems into the sustainability report [INT5]. Considering the amount of work, Interviewee 1 gave the example of Shippeo where it is laborious to connect all the right sub-contractors within the platform [INT1]. Interviewee 4 added that "*we also learned that it's terribly labour-intensive to be able to achieve all that data with the right percentage of deviation that you allow*" [INT4]).

4.4 DISCUSSION OF INTERVIEW RESULTS

Section 4.3 focused on the consolidation of the interview results, aiming to answer, *"What are the practical challenges and experiences that companies encounter while implementing CSRD reporting, particularly in terms of motivational drivers and technical boundaries?".* To help answer this main research question, three additional sub-research questions were formulated. Therefore, these will now be resolved in this discussion.

4.4.1 Motivation

To create an understanding of the motivation of companies regarding CSRD reporting and digitalization, RQ 4 investigates "*What are the motivational drivers for companies regarding software adaptation and digitalization in the context of CSRD reporting?"*. To answer this question, we first examined the extent of motivation and organizations' motivational drivers concerning CSRD reporting in general (RQ 4A: *To what extent are companies motivated to commit to CSRD reporting and what are the drivers behind this motivation?*).

The findings indicate a range of motivations among organizations concerning CSRD reporting. Three organizations experience that stakeholders are inclined to fulfil only the minimum necessary requirements, not wanting to go the extra mile. Conversely, another organization perceives CSRD as a compliance exercise, stating they have already engaged in more extensive sustainability reporting than the CSRD requires. This aligns with the attitude of companies part of Eklund & Vaaler's research, who indicated that their motivation is mainly from a regulatory perspective [54]. Nonetheless, a notable group of five organizations is motivated to exceed the CSRD requirements, with two proactively pursuing compliance voluntarily out of a commitment to sustainability.

Examining the underlying drivers of these motivations, five organizations aspire to be leaders in sustainability, with one even taking on a central position in facilitating the reporting process for their stakeholders. Furthermore, they aim to introduce new perspectives and highlight their successes. The potential to create business value is another significant driver, as insights derived from CSRD reporting can lead to cost savings, innovations, and reduced interest rates for loans. However, for organizations with less motivation, general resistance to policy is experienced, as well as difficulties in aligning the CSRD with their existing sustainable practices. The opportunities are in line with findings from Glaveli et al. [59], but their insight that sustainability is perceived as a threat to the business was not expressed by any of the interviewees in this research.

Additionally, sub-RQ 4B aims to find "*How do the choices of companies for strategies and tools reflect their motivation and commitment to digitalization?".* Overall, companies' choices reveal a strong motivation and commitment to digitalization, primarily driven by the necessity for enhanced sustainability reporting and operational efficiency. For instance, organizations strive for a centralized digital landscape including various useful tools, demonstrating a clear commitment to integrating comprehensive digital solutions to support sustainability. However, this commitment varies, with six organizations managing digitalization in-house while others outsource these efforts entirely.

Four organizations currently employ tools for collaborative reporting and emission calculations, indicating an advanced commitment to leveraging digital solutions for ESG data management. Three other organizations adopt a balanced approach, anticipating future integration of comprehensive systems, while currently using a combination of tools and manual methods. This showcases a more pragmatic digital strategy, aligned with immediate operational needs.

Similarly, one other organization is more cautious in their digital approach, initially relying on Excel (just as indicated by [13]), despite auditors marking this as unreliable. Nevertheless, they expressed intentions to evolve digitally in the future. Considering that research by Atanasov highlighted that digital transformation can enhance sustainability reporting [2], it is good that organizations have this commitment, even though some may fully embark on this process only at a later stage. All of the interviewed companies do understand the need to transform from analogue reporting to digital and automated processes, as discussed in [12], [54].

Examples of software solutions to calculate emissions are Big Mile, Milieubarometer, the platform by Realised, and Coolset's tool. Additionally, Shippeo is used for transport visibility. Considering report writing, Workiva and Nossa Data are currently utilized by organizations. Five companies combine the use of these tools with a data warehouse or data lake or plan to implement one in the future. In general, the future of these systems is expected to evolve significantly as the field matures, which is why organizations tend to be careful with early commitment to novel solutions. This hesitation towards new software solutions was also highlighted by de Vries [6]. Considering that only every fifth organization in the research by PwC planned to use a dedicated software solution [13], the interview findings show that this is currently about three quarters.

Overall, the varying degree of digital tool adoption across these organizations underscores a universal acknowledgment of digitalization's critical role in achieving sustainability goals. The integration of tailored digital solutions reflects their commitment to transparency, efficiency, and strategic alignment with broader sustainability objectives.

4.4.2 Implementation process

After discussing companies' motivation for CSRD reporting, RQ 5 aims to understand what the implementation process has been for companies so far and what their future milestones are (*"What is the implementation process that companies are going through for CSRD reporting?"*).

Given the timeframe for CSRD implementation, four companies initiated the process about two to three years ago. Others started more recently, commencing at the beginning of 2023 or even 2024. Regardless of the initiation period, all interviewed organizations that fall within the mandatory reporting group underscored the necessity of starting on time, emphasizing the substantial workload and the tedious nature of the process.

To lead the CSRD implementation, each interviewed organization established a dedicated CSRD team. Nevertheless, the composition of these teams varies significantly across the organizations. In two cases, sustainability departments are responsible for overseeing the process, whereas four other companies have formed working groups comprising three to six employees from various departments. In larger organizations, these working groups can include up to 15 people, extending to approximately 150 individuals directly involved across the organization. Another organization positioned the process within the finance department, whilst additionally including employees from other departments.

The actual processes led by these teams vary as well. Three companies began by fostering support within the organization, recognizing the involvement of numerous internal stakeholders. It is interesting to note that while research by Greiling & Bauer considers forming multidisciplinary teams to be complicated [53], this was not reflected in the responses of the interviewees in the present study. As a next step, four companies initially conducted their own internal materiality assessment and are now in the process of the official DMA. Other than this taking significant time to complete, no specific challenges were emphasized compared to the challenges highlighted in [55]. Simultaneously, five organizations already work on the technical infrastructure and requirements regarding the data points. To aid them with this process, two organizations made enterprise architecture models to provide an overview of their IT landscape. Others (additionally) created an overview of each data point, tracking various aspects ranging from definitions to source systems.

Future milestones include finalizing the DMA before the end of the year to clarify the material themes. Five organizations additionally aim to improve their digital landscape but plan to postpone this for approximately a year to better understand forthcoming developments. A significant factor influencing this delay is the evolving regulatory landscape, which creates uncertainties and challenges in committing to specific actions. Another anticipated challenge is the collection of data across the entire value chain.

Throughout this implementation trajectory, seven organizations have found the process to be complex. Concerns have been expressed about the vagueness, complexity and novelty of the CSRD, as was found by de Vries in 2023 [6]. Together with the extensive scope that impacts numerous stakeholders, the process is being experienced as difficult to navigate and interpret. Besides this complexity, challenges such as a lack of capacity for CSRD implementation and insufficient central direction have been identified, which was apparent in many previous studies [6], [7], [12], [51], [52]. Another critical aspect raised in this context is the need for incentives and creating an engaging process for internal stakeholders.

Alongside these challenges, four interviewees highlighted the importance of transparency. During the implementation process, organizations should be honest about the origins of reductions in emissions and the motivators behind their decisions. These insights from practice align with the findings of Eklund & Vaaler [54].

A final significant aspect of the implementation process is the inclusion of external stakeholders, both as part of the value chain or as advisory parties. Advisory parties provide support to organizations in achieving CSRD compliance and in selecting appropriate tools. Leveraging their experience from working with various companies, these advisors can offer valuable insights. Regarding stakeholders within the value chain, there are varying levels of cooperation, motivation, and data availability. Company H confirmed the identified challenge of engaging with stakeholders outside of Europe [52], [55], noting they tend to lack the feeling of necessity.

It is noteworthy that none of the interviewees addressed the topic of digital tagging (explained on page 32), during discussions concerning their implementation process, future milestones, or any other relevant topic. This may suggest that the digital tagging requirement is intended to be addressed at a later phase of the implementation process or it can indicate a lack of consideration or awareness among companies. Therefore, this observation can serve as a potential avenue for future research.

In summary, the implementation process for CSRD reporting involves establishing dedicated teams spanning multiple departments, conducting materiality assessments, and developing technical infrastructure. Challenges include regulatory uncertainties, capacity constraints, and complexity. Organizations emphasize the importance of timely action, transparency, and stakeholder engagement, with varied approaches across organizations. Despite certain difficulties, progress is currently being made towards compliance.

4.4.3 Technical considerations

To get an understanding of the technical boundaries for organizations, RQ 6 addresses "*What are the technical constraints companies face in achieving CSRD reporting compliance?*". This inquiry considers the metrics companies report on, the sources of this data, the extent of data availability, and additional challenges that are encountered.

Considering sub-RQ 6A, "*What metrics do companies report on in a CSRD context?*", companies report on a range of metrics associated with ESG themes, each prioritizing their data points based on the outcome of the DMA. Climate change and emissions are universally critical, and much effort is made to gather data. This is an interesting outcome, as the research by Bauer & Greiling showed that the environmental category was initially more part of long-term plans instead of immediate action [58]. Additionally, themes such as pollution, resource use, circular economy, workforce conditions, workers in the value chain, consumer impacts, and business conduct are covered. Specific metrics that are reported on include emissions, consumption, material use, kilometres driven, waste, turnover, and additional employee metrics. Smaller companies tend to begin with a limited scope of themes and plan to expand as their data management strategies evolve. Larger organizations may have multiple KPIs in place, however, these often do not encompass all material themes yet.

Sub-RQ 6B then aims to find "*From which sources is the required data obtained?*". Data for these metrics is sourced from various systems. Primarily, companies utilize ERP systems, such as AFAS and SAP, where much of the data is stored. Additionally, data collection often involves manual processes, utilizing supplier invoices, annual statements, and HR systems. Oftentimes, Excel sheets are employed either as data source or as calculation tool. Three organizations have integrated analytical tools like Power BI to derive numerical insights. In cases where ERP systems are not utilized, data is gathered from a few major applications. Data exchange with stakeholders is often performed via email, using PDFs and Excel files.

Moving on, the focus shifts to data availability for sub-RQ 6C (*"To what extent is the necessary data available for companies?").* The availability of the necessary data significantly varies among organizations; while five companies possess comprehensive datasets, others consider having about 50 percent available. The data is often dispersed across various sources, complicating consolidation efforts. One organization acknowledged that most of their available data is financial, having a lack of operational insights.

Apart from the data availability, several additional challenges were identified for sub-RQ 6D ("*What are challenges organizations encounter concerning data?"*). First, organizations face difficulties due to the varying formats and levels of detail in which stakeholders provide data. These discrepancies were initially identified in research by [57]. Second, specific software solutions require data in particular formats, needing manual adjustments. Furthermore, the core functionalities of many systems are not designed for reporting purposes, coming with misalignments in data hierarchies. Translating information into usable data also presents significant challenges. Data quality emerged as a primary concern, with error detection being complicated by estimations and inconsistent stakeholder data. This is closely related to incomplete data, which necessitates assumptions and estimations, further complicating the process. Lastly, manual data handling introduces additional risks, being both error-prone and labour-intensive. This concern was also present for Norwegian companies in earlier research [54].

All in all, organizations face several technical constraints in achieving CSRD reporting compliance, including data fragmentation across multiple systems, reliance on manual data collection, inconsistent data formats, and inadequate system functionalities for reporting. Data quality and availability also present challenges, with several companies lacking comprehensive data and struggling with error detection due to incomplete or estimated data. Summarizing these aspects, the main challenge indeed comes down to consolidating all data into a system for reporting purposes, as expressed by de Vries [6] and Matilla & Sasi [52].

4.5 CONCLUSION ON PRACTICAL CHALLENGES AND EXPERIENCES OF CSRD

Finally, the answer to the main research question of this chapter consolidates all relevant aspects into a comprehensive overview:

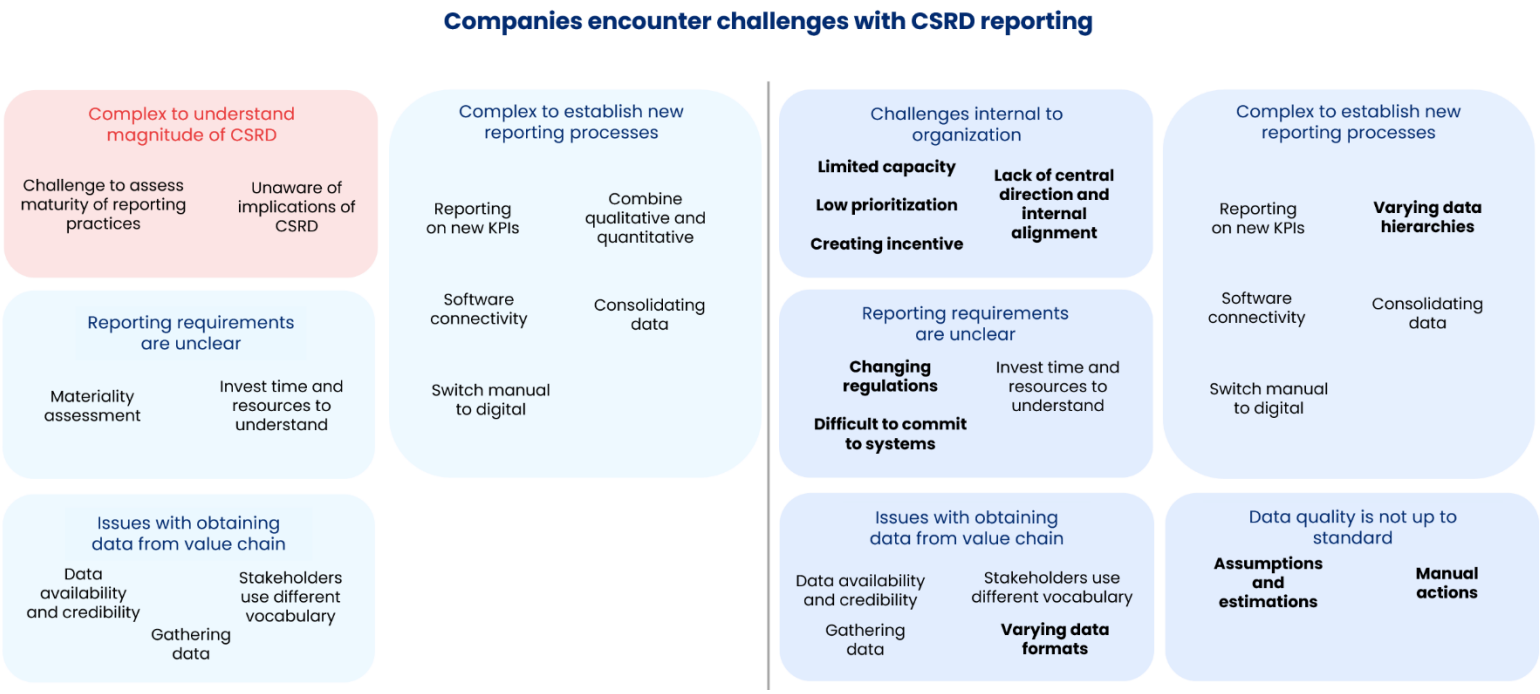
> *"What are the practical challenges and experiences that companies encounter while implementing CSRD reporting, particularly in terms of motivational drivers and technical boundaries?"*

Insights from the interviews reveal that organizations face several practical challenges and varied experiences during the CSRD implementation process. Motivational drivers vary widely, with some companies viewing CSRD reporting as a compliance necessity while others aim to exceed the requirements out of their commitment to sustainability. Key motivations include the desire to lead in sustainability and enhance business value. Despite this, resistance to policy and difficulties in aligning CSRD with existing practices persist in some organizations.

Technically, the implementation process is complex, often involving dedicated CSRD teams from various departments and extensive internal and external stakeholder engagement. Initiated about two to three years ago by most organizations, the process entails the double materiality assessment, developing technical infrastructure, and overcoming regulatory uncertainties. Common challenges include the vagueness and novelty of CSRD requirements, capacity constraints, and the necessity for transparency and incentivised stakeholder engagement.

Technical constraints are prominent, with companies reporting on a range of ESG metrics sourced from diverse systems as well as through manual data collection from supplier invoices and annual statements. Integration of analytical tools like Power BI is sometimes done, yet data fragmentation, varying data formats, and inadequate system functionalities pose significant hurdles. Data quality and availability are inconsistent, often leading to error-prone and labour-intensive processes. The evolving regulatory landscape further complicates early commitment to digital solutions, with organizations carefully navigating these complexities to achieve compliance.

Overall, organizations are making progress toward CSRD compliance, despite encountering significant motivational and technical challenges that require strategic, technical, and organizational adaptations. Figure 16 compares the challenges identified in literature (Figure 1), presented on the left side, with those emerging from the interviews, depicted on the right side. Novel aspects identified through the interviews are emphasized in **bold**.



Red indicates a category identified in literature that was not applicable in practice.

Bold challenges are novel challenges identified in practice compared to existing literature.

Figure 16 - Updated overview of challenges

A comparative analysis reveals that the companies participating in this study have developed a comprehensive understanding of the CSRD's magnitude; this presents a shift from earlier findings (indicated in red). Although the reporting requirements remain unclear, there has been a notable change: materiality assessments are no longer perceived as a significant challenge, whereas adapting to evolving regulations has emerged as a new concern. Challenges related to obtaining data from the value chain persist, with the additional complexity of varying data formats. Within the theme of establishing new reporting processes, the issue of varying data hierarchies has been introduced. Moreover, the interviews revealed a new category of challenges concerning the quality of data, due to assumptions, estimations, and manual actions.

Altogether, these insights serve as the foundational elements for the development of the roadmap. By examining the challenges identified through empirical findings, we can ensure that the roadmap is tailored to address the specific obstacles faced by organizations in their journey towards CSRD compliance. In Chapter 5, the focus shifts to the design of the roadmap, where these challenges will be accounted for when drafting the guidelines.

5 ROADMAP DESIGN

In this chapter, we combine insights from literature and the semi-structured interviews into a roadmap for companies to guide the process of data consolidation for CSRD reporting. This is the main goal of the research, and this chapter describes the design of the treatment for achieving this goal. Section 5.1 discusses the purpose and scope of the roadmap, followed by the design methodology in Section 5.2. Subsequently, the six phases of the roadmap are outlined in Section 5.3 with the actual roadmap in Section 5.4. The answer to the main research question is presented in Section 5.5, through a discussion of the roadmap.

5.1 PURPOSE AND SCOPE

The purpose of the roadmap is to provide a structured approach for data consolidation in the context of CSRD reporting. Specifically, the roadmap addresses the main research question:

> *"How can companies integrate data from various sources within their IT architecture and along their value chain, facilitating data consolidation for CSRD reporting?"*.

One of the steps under the treatment design in the DSRM is to consider available treatments [15]. While several existing frameworks, such as the roadmap developed by BDO [106], provide valuable guidance towards CSRD compliance (as illustrated in Figure 17), they fall short in offering specific steps for achieving "integrated reporting" during the reporting phase. To the best of our knowledge, existing roadmaps lack concrete guidelines on how to operationalize data integration for CSRD purposes. Consequently, we focus on developing a roadmap that fills this gap, providing detailed steps that organizations can implement to achieve integrated reporting. This roadmap is designed to be utilized by both companies directly as well as consultancy firms, such as Flawless Workflow, that support their clients in navigating the process of CSRD implementation.

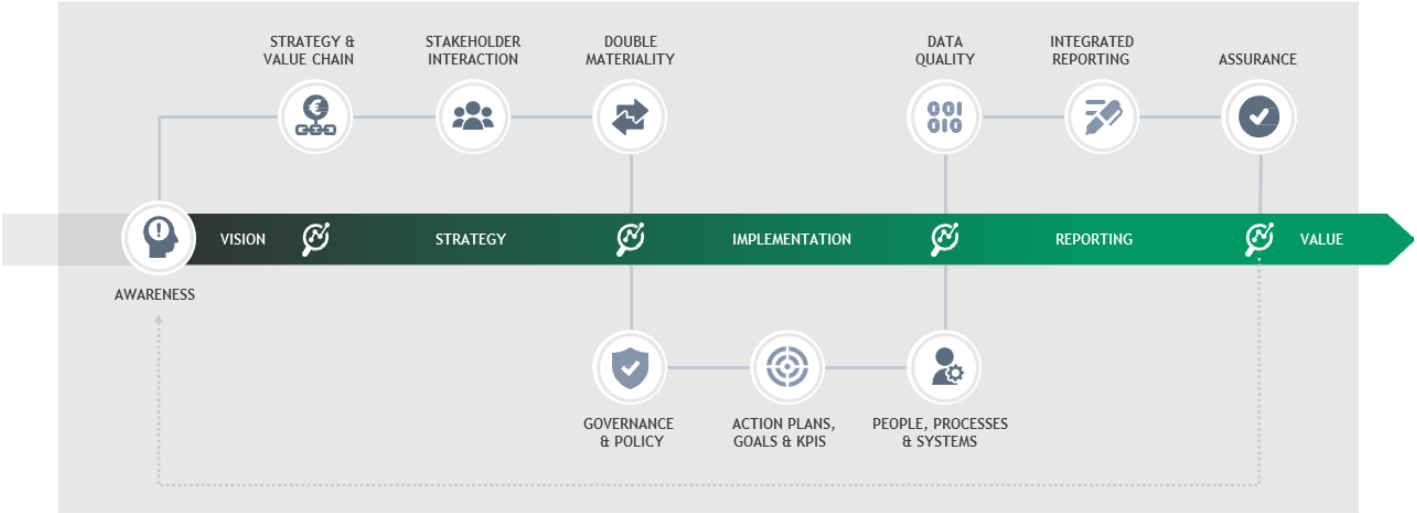


Figure 17 - Roadmap for CSRD compliance by BDO (2024)

The roadmap developed in this research requires to incorporate concrete steps towards integrated CSRD reporting. Moreover, it must address the specific challenges identified during the interviews conducted as part of this study.

5.2 DESIGN METHODOLOGY

In constructing a roadmap, the DSRM is used more often as a general framework [16]. Despite an extensive review of existing literature, we have not been able to identify any specific design methods applicable to the context of roadmap development. Therefore, the following five steps are derived from the Design Process (Appendix 9.9) [107], with the original steps indicated in *italic*:

1. **Identification of best practices and challenges** – *define the problem & collect information*: the first step involves a thorough analysis of the interview results, focusing on the identification of best practices among companies as well as the challenges they face in achieving integrated reporting.
2. **Assessment of literature review** – *collect information*: the second step focuses on the insights obtained with the literature review, extracting relevant insights that inform the roadmap. This includes identifying both challenges and potential solutions from existing research.
3. **Solution development** – *brainstorm and analyse & develop solutions*: based on the identified challenges, solutions are sought through an examination of the literature review findings. Where necessary, additional research is conducted to address gaps in the existing knowledge.
4. **Categorization of phases** – *develop solutions*: the fourth step involves organizing the empirical and literature-based insights into six main phases.
5. **Roadmap construction** – *develop solutions*: finally, the roadmap is constructed according to a structure that aligns with the standard roadmap model of Flawless Workflow. Each phase in the roadmap is detailed with excerpts from the empirical data and literature, offering a clear path for companies to achieve integrated CSRD reporting.

In the original Design Process, Step 5 and 6 are the presentation of ideas to others to collect feedback and improving the design, respectively [107]. These steps will be discussed in Chapter 6 as they are part of the DSRM treatment validation. Altogether, this approach ensures that the developed roadmap is both theoretically grounded and practically relevant, addressing the specific needs and challenges identified through empirical research.

5.3 ROADMAP PHASES

This section aligns with step four of the design methodology as described in Section 5.2. Based on existing roadmaps created and utilized by Flawless Workflow, six phases were created. These phases are (1) creating support, (2) landscape mapping, (3) overview of data points, (4) solution research, (5) centralizing data, and (6) to-be landscape realization. For each phase, the empirical and literature-based insights are discussed. Additionally, each section finishes with a summarizing excerpt which is incorporated in the roadmap.

5.3.1 Phase 1 – Building support

The first phase of the roadmap focuses on fostering support and commitment for CSRD reporting, engaging both internal and external stakeholders. To ensure employees grasp the significance of CSRD reporting and commit to its quality, it is essential to foster support, often referred to as *"draagvlak"*. Establishing a sense of urgency is a key strategy in achieving this, which can be accomplished by actively involving employees in sustainability initiatives and embedding them in the reporting process. Such engagement can enhance their awareness of the importance of CSRD, thereby fostering a collective urgency [INT2]. Additionally, it is important to clearly communicate the value of CSRD reporting and how it offers new perspectives for accelerating ESG goals [INT1, INT7]. Employees must understand the data that is being entered, as the accuracy and quality of the report are directly dependent upon the quality of the data input [INT1, INT8].

Beyond general engagement, it is recommended to assemble a diverse group of employees, representing various departments within the organization, to take responsibility for the reporting process. This cross-departmental collaboration ensures a comprehensive and multifaceted approach to CSRD reporting [INT4, INT6]. It is recognized, however, that transparency and reporting are not often perceived as something fun [INT7], given that they necessitate changes to established working methods. To effectively manage these changes, it is important to implement change management practices within the organization.

According to Oakland & Tanner, successful change management involves aligning the need for change with operational issues, thereby ensuring that all organizational members understand how the change will impact them and what will be required of them [108]. Senior management must demonstrate commitment by taking a leadership role in this transformation. Moreover, process thinking should be emphasized to ensure that organizational processes are fully understood, systematically measured, and continuously improved. Oakland & Tanner also highlight the added value of external consultancy in facilitating change and transferring knowledge. To support these behavioural changes, organizational culture must be aligned with the new requirements. Lastly, continuous review mechanisms are necessary to monitor and adapt the change approach as needed [108].

Additionally, it is essential to establish a central direction and foster internal alignment [INT4, INT5], so that all members of the organization work towards a common goal.

Translating the CSRD requirements into the specific context of each department can facilitate better understanding and alignment [INT6].

In terms of external stakeholders, challenges related to transparency are prevalent as well [INT1]. Therefore, initiating conversations with stakeholders about CSRD requirements and data exchange is critical. For Company A, for instance, it became evident that their stakeholders expect them to lead these efforts [INT1]. It is therefore important to recognize what position the company has in the value chain. Considering the difficulties associated with varying data vocabulary, as identified in both the literature and practice, it is advisable to work towards establishing a standard for terminology and data exchange in collaboration with key stakeholders to address this issue effectively.

Phase 1 – Building support

- **Engage and involve employees**: actively involve employees in sustainability initiatives and the CSRD reporting process to create a sense of urgency and commitment to quality reporting.
- **Communicate the value of CSRD**: clearly articulate how CSRD reporting contributes to achieving ESG goals and offers new perspectives.
- **Establish a cross-departmental reporting team**: form a diverse team of employees from different departments to take responsibility for the reporting process, fostering collaboration and a broad coverage.
- **Implement effective change management practices:** align the need for change with operational realities, involve senior management as change leaders, and ensure continuous review and adaptation of the change approach.
- **Standardize exchange with stakeholders**: initiate discussions with external stakeholders to address challenges related to varying terminology and work towards establishing a standard for consistent data exchange.

5.3.2 Phase 2 – Mapping the IT landscape

The second phase of the roadmap focuses on mapping the IT landscape, a critical process in understanding the current state of an organization's technical infrastructure. The first step involves visualizing the existing IT landscape to establish a foundational understanding of the source systems, following the approach taken by Companies A and G. This visualization provides a base-level understanding of the connectivity between the systems in use and their interactions, which is essential for identifying and addressing issues such as varying data hierarchies. While it is possible at this stage to begin mapping potential future states of the IT landscape, this task can also be suspended until after the completion of the solution research in Phase 4.

Enterprise Architecture facilitates the visualization of the interaction between business processes and IT systems [65]. As noted in the literature, EA provides a comprehensive overview that is valuable for stakeholders, enabling organizations to optimize business activities, refine strategies, and more effectively leverage IT resources [67]. In the context of CSRD, EA is particularly beneficial as it clarifies the relationships within the IT landscape, mapping out data flows and interactions [INT7]. This allows organizations to gain a

thorough understanding of their data sources, the connections between various applications, and potential pathways for integration.

Moreover, it is important to extend this mapping exercise beyond the internal IT landscape to include interactions with external stakeholders, particularly considering data sharing along the value chain. By incorporating these external connections, organizations can develop a more comprehensive view of their IT environment and how external ESG data is retrieved from stakeholders.

Given that TOGAF and ArchiMate are the most prominent EA framework and EA modelling language, respectively, it is recommended that organizations utilize a combination of these methodologies. Additionally, the literature review showcased that ArchiMate has several features important for modelling the IT landscape in a CSRD context (Table 4). This approach enables a comprehensive visualization of the IT landscape, involving perspectives from business, application, and technology domains. Therefore, it provides a robust foundation for informed decision-making on CSRD matters. However, if organizations do not have the necessary skills in-house, it can be decided to create a simpler visualization of the landscape such as a more simple flow chart.

To provide a conceptual understanding of potential EA models, Figure 18 and Figure 19 have been designed using ArchiMate as language and Archi as tool, illustrating both an as-is and a to-be situation with a hypothetical IT landscape for a sample company. The as-is

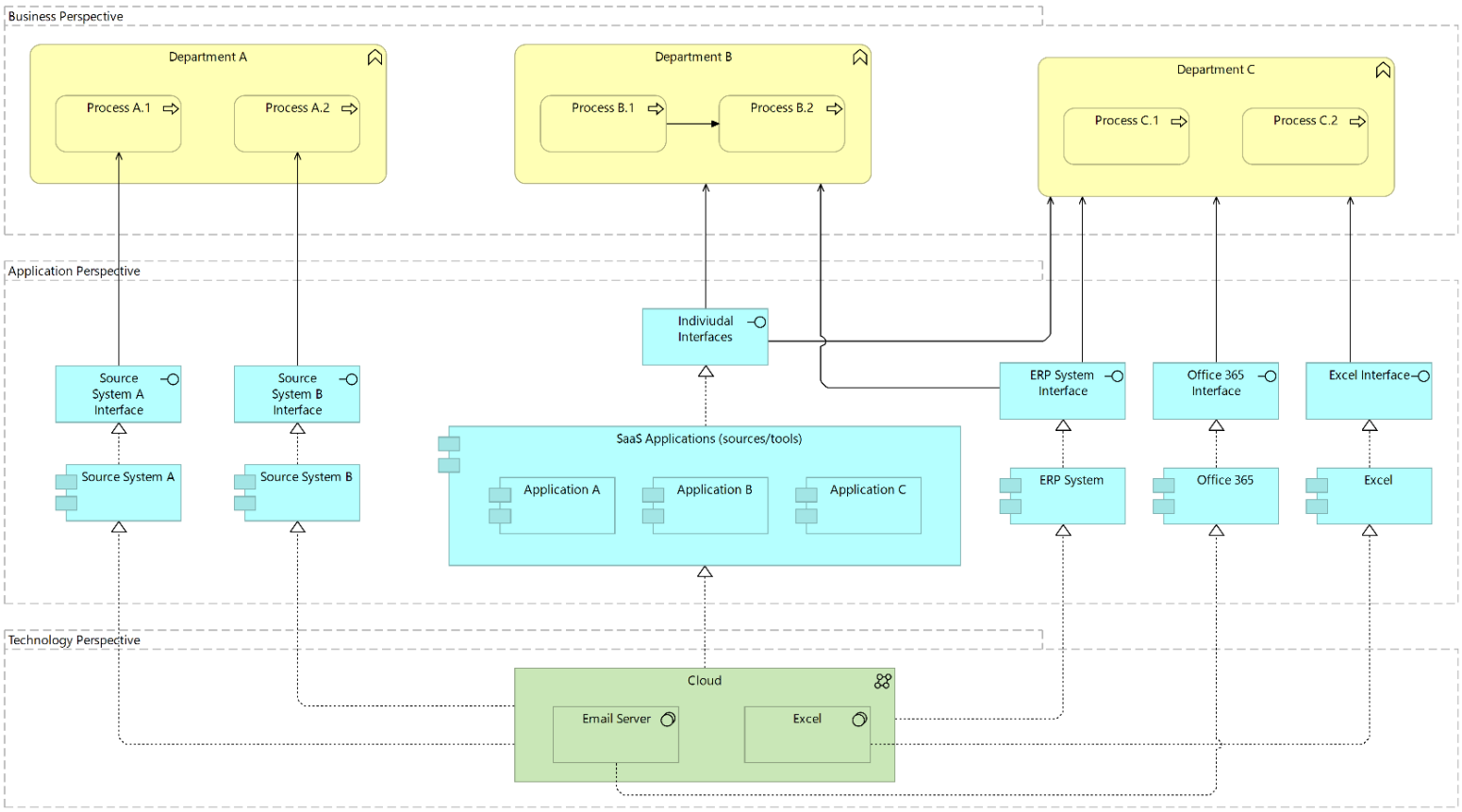


Figure 18 - Fictitious as-is EA model of an IT landscape

model depicted in Figure 18 represents the current state of an organization's business architecture, including cross-layered relationships. These departments in turn utilize different applications and source systems, supported by a cloud platform. Three distinct departments operate with independent processes, which are in turn supported by a variety of applications and source systems. All of these are supported by a cloud platform in the technology layer.

In exploring potential improvements to the IT landscape to facilitate CSRD reporting, Figure 19 presents the to-be model that reflects a more integrated approach. In this future architecture, stakeholders across the value chain are actively engaged, contributing and gathering data alongside the core organization. A business collaboration highlights the importance of cooperation for sustainability reporting. The model additionally introduces new processes, such as the creation of standardized Excel and PDF templates by the core organization, which are created to facilitate data sharing among stakeholders.

At the application layer, the introduction of a data consolidation platform, such as a data warehouse, lake, or lakehouse, serves as a central hub for data integration (indicated in Figure 19 with an orange outline). All applications, including both source systems and CSRD

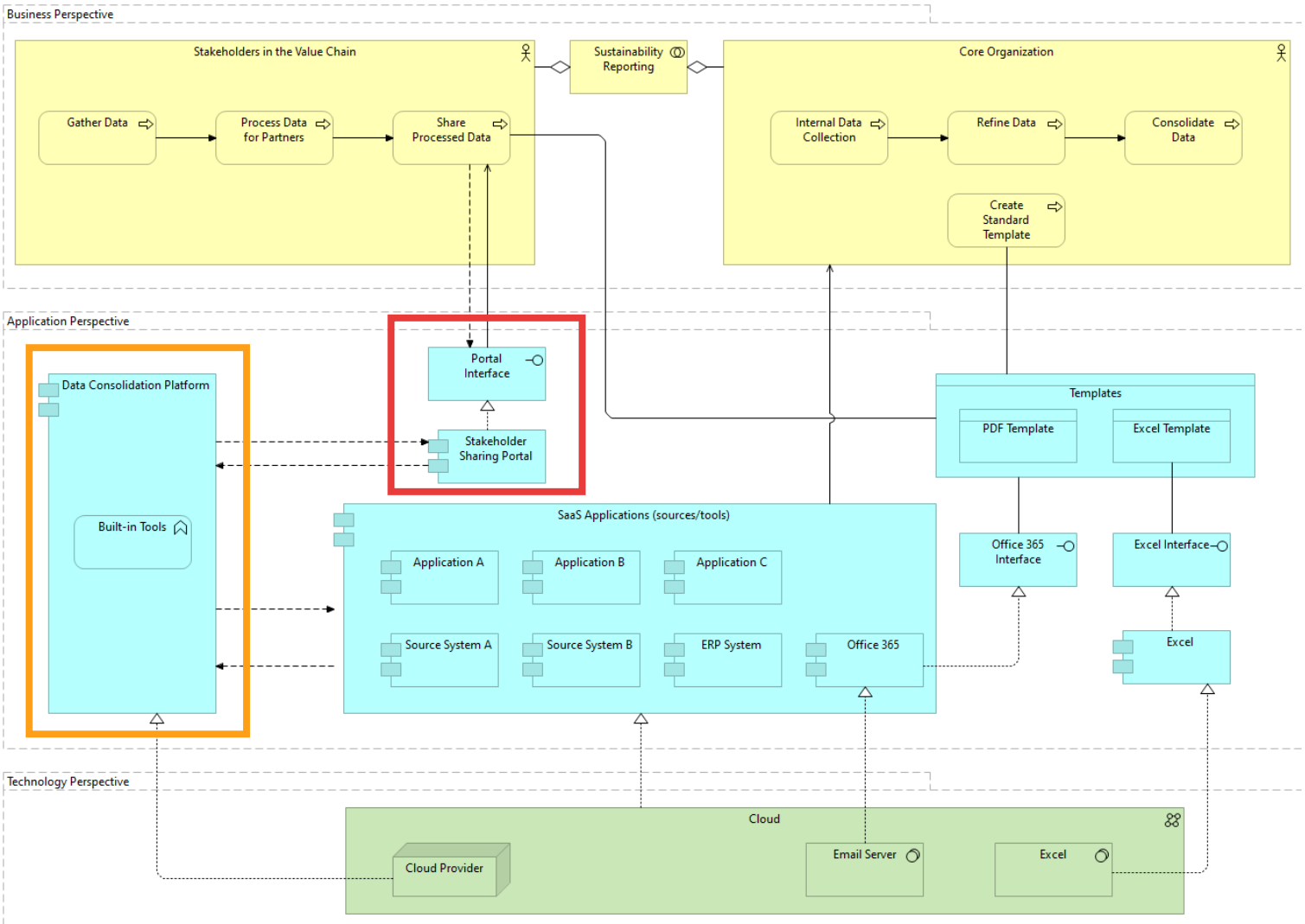


Figure 19 - Fictitious to-be EA model of an IT landscape

tools, are interconnected with this platform, enabling continuous data exchange. If it is impossible to connect certain source systems to the consolidation platform, essential data could be exported in a PDF or Excel format. Additionally, a data-sharing portal (red outline) has been established, providing stakeholders with an interface for inputting their data. This portal could also offer potential functionalities for stakeholders to access data from the core organization. As in the as-is model, the technology layer continues to be supported by a cloud platform. These ArchiMate models can be designed with tools such as Archi[15] or Bizzdesign Enterprise Studio[16].

Phase 2 – Mapping the IT landscape

- **Visualize the current IT landscape**: create a comprehensive visualization of the existing IT landscape to establish a clear understanding of source systems and their interactions. Utilize EA practices to map relationships, visualize data flows, and provide a comprehensive overview that supports business-IT alignment.
- **Use TOGAF and ArchiMate**: apply a combination of TOGAF and ArchiMate to achieve a comprehensive visualization of the IT landscape from business, application, and technology perspectives.
- **Incorporate external stakeholders in mapping**: extend the mapping process to include interactions with external stakeholders, particularly in the context of data sharing along the value chain.
- **Assess potential issues**: evaluate the connections between applications and identify potential issues, such as inconsistencies in data hierarchies, to ensure seamless integration.
- **Visualize potential future situations**: based on necessary changes for CSRD reporting, potential future versions of the landscape can be visualized. After assessing the options, the selected scenario will serve as the blueprint guiding the IT development efforts.

5.3.3 Phase 3 – Overview of metrics

Upon achieving an understanding of the IT landscape, including identification of CSRD source systems, Phase 3 shifts the focus to the systematic identification and assessment of the required data points. The first step of this phase is to generate an overview of the metrics necessary for CSRD reporting, guided by the outcomes of the materiality assessment [INT3]. The next step involves establishing a clear "Definition of Done" for each data point [INT7], which will encompass the evaluation of their availability, source system, and reliability, as well as identifying areas for improvement in these domains [INT3, INT6]. This ensures that for each data point, certain aspects are systematically checked.

[15] Archi is a free and open-source visual-modelling and design tool for creating ArchiMate models and modelling sketches.

[16] Bizzdesign is a Dutch enterprise architecture and BPM SaaS platform vendor, known for the co-development of ArchiMate and the development Bizzdesign Enterprise Studio.

For each identified metric, it is crucial to conduct a gap analysis based on the Definition of Done, to determine aspects such as missing data and to outline strategies for addressing any inconsistencies [INT3]. Additionally, this phase should anticipate and plan for future data quality improvement initiatives to ensure ongoing compliance and accuracy in reporting [INT1]. Table 7 presents an example for the overview of metrics. Suggested measurements are offered for each attribute, though alternative methods can be applied. Moreover, additional attributes can be incorporated as appropriate for each organization.

Table 7 - Suggestion for the overview of metrics, based on data points of ESRS E1 – Climate change

Metric	Data type	Source system	Reliability	Complete
Financial resources allocated to action plan	monetary	System A	10/10	100%
Achieved GHG emission reductions	ghgEmissions	System B & C	6/10	50%
Total energy consumption related to own operations	energy	System B	8/10	80%
Percentage of renewable sources in total energy consumption	percentage	System B	7/10	90%
Net revenue from customers operating in oil-related activities	monetary	System A & D	4/10	40%

Furthermore, this phase can involve another mapping exercise, specifically focusing on the visualization of data process flows [INT7]. These process flows can be depicted either through additional models or as extensions of the as-is IT landscape mapping developed in Phase 2. These models include the origins of data, the systems in which the data is utilized, and potentially even the personnel who have access to it, thereby providing a detailed understanding of the data's lifecycle within the organization [INT7].

Phase 3 – Overview of metrics

- **Identify required metrics**: generate a comprehensive overview of the metrics necessary for CSRD reporting, informed by the materiality assessment.
- **Establish a "Definition of Done"**: define criteria based on which each metric will be assessed, including for instance availability of data, the source system, reliability, and potential areas for improvement.
- **Conduct a gap analysis**: assess each data point according to the definition, to identify gaps between current data availability and reporting requirements and develop strategies to address these gaps.
- **Plan for future data quality improvements**: identify and plan for ongoing efforts to enhance data quality and ensure sustained accuracy in reporting.
- **Visualize data process flows**: map out the data process flows to understand the origins, usage, and access points of data within the organization, either as additional models or extensions of the existing IT landscape mapping.

5.3.4 Phase 4 – Solution research

With a clear understanding of the data points and reporting requirements, Phase 4 is dedicated to solution research on available tools for reporting and CSRD data consolidation. This research should begin by defining a set of criteria against which potential solutions are evaluated, which may include the development of user stories to simulate practical applications. During the interviews, several tools were identified as potential solutions, including an LCA programme [INT2], Milieubarometer [INT3], Big Mile, Shippeo [INT1], the platform by Realised [INT4], Workiva [INT5], Nossa Data [INT7], the platform by Coolset [INT8], and Power BI [INT4, INT8]. Furthermore, an important aspect to consider during this solution research is the capability for digital tagging and how that can be facilitated by certain tools. Another aspect is the possibility for these tools to be connected to a data warehouse or data lake.

Focusing on the Power BI tool, there is generally a difference between standardized dashboarding solutions and a fully custom Power BI. As various financial templates are already available, it is likely that similar templates will be developed for CSRD purposes. Nonetheless, as with other existing templates, these pre-designed solutions may not fully address the specific needs of every organization. Consequently, it may be more advantageous for some organizations to develop a tailored Power BI solution to better align with their unique reporting and analytical requirements.

One challenge highlighted in the interviews is the manual nature of data collection processes, where numbers are often manually extracted from PDFs [INT3, INT4, INT8]. Therefore, Optical Character Recognition (OCR) technologies can be explored, which are designed to extract textual content from digital images and converted PDF files [109]. However, it is important to note that traditional OCR tools are typically optimized for documents in straightforward narrative formats and may have issues with more complex layouts, such as those found in invoices and annual statements. Given the ongoing advancements in OCR technology, continuous monitoring of developments in this area is recommended.

As the final step of Phase 4, the model of the future IT landscape should be updated to incorporate the newly selected solutions, ensuring that the to-be scenario reflects the latest technological decisions. If a to-be model has not yet been developed, it should be created at this stage.

Phase 4 – Solution research

- **Define evaluation criteria for solutions**: establish clear requirements and develop user stories to guide the assessment of potential tools for CSRD reporting.
- **Explore available tools**: perform market research to investigate various tools, taking the established criteria into consideration.
- **Address manual data collection challenges**: research the applicability of OCR technologies to streamline data extraction from complex documents, while staying informed on advancements.
- **Create or update the to-be model**: incorporate the selected tools into the to-be IT landscape model, ensuring it accurately reflects the chosen technological solutions for CSRD compliance.

5.3.5 Phase 5 – Centralizing data

Once the appropriate software and tools have been selected, the next step is the centralization of data. This can be effectively achieved through the implementation of a data warehouse, data lake, or the recently introduced data lakehouse. Based on challenges identified in both the interviews and the existing literature, each of these techniques serves as a viable solution for centralizing data spread over several sources [INT4].

Additionally, concerns related to evolving regulatory requirements have been noted [INT1, INT2, INT5]. Consolidation tools provide the flexibility necessary to connect diverse applications and facilitate transitions whenever needed. Notably, organizations can initiate reporting efforts with tools such as Power BI without making a commitment to any specific CSRD reporting tool.

Moreover, organizations strive for a single source of truth [INT1, INT2], and consolidation tools present significant opportunities to achieve this objective. For instance, while ERP systems typically house most of the relevant data, the incorporation of external data sources—such as information from corporate groups [INT8] or external stakeholders [INT1]—enhances the overall quality and reliability of the data. This integration mitigates the risks associated with traditional data-sharing methods, such as email, thereby ensuring that data integrity is maintained throughout the reporting process [INT7].

As discussed previously (Sub-section 3.2.3), for each of the techniques different use cases can be presented. Data warehouses are particularly suitable for smaller organizations with a more streamlined value chain and structured data needs. Designed primarily for reporting and business intelligence, data warehouses offer user-friendly interfaces that do not require extensive data science expertise.

Data lakes, by contrast, are more appropriate for larger organizations with extensive value chains and a need to manage both structured and unstructured data. The scalability and flexibility of data lakes support a wide range of use cases, from basic reporting to advanced financial predictions. However, they require robust governance and significant data science expertise to avoid becoming disorganized "data swamps". The potential to leverage machine learning and artificial intelligence makes data lakes ideal for organizations with the resources to invest in these technologies.

Data lakehouses represent a hybrid solution, combining the strengths of both data warehouses and data lakes. Suitable for larger organizations with the resources and ambition to adopt novel technologies, data lakehouses support the storage of both structured and unstructured data, offer user-friendly interfaces similar to data warehouses, and accommodate a wide array of use cases. Despite their advantages, data lakehouses are relatively new and demand skilled technical resources, which may be a limiting factor for some organizations. However, for those willing to invest in this innovative approach, data lakehouses offer significant potential for efficient data consolidation and advanced analytics in the CSRD reporting process.

All in all, the choice between these techniques should be guided by the specific needs and capabilities of the organization. Regardless of the chosen approach, preparation of data and systems is essential; it addresses several issues highlighted during the interviews. One concern that was raised is the inconsistent calculation formats employed by both internal and external stakeholders for the same data [INT1, INT7]. To address this challenge, an ETL process can be implemented to standardize data [110]. This begins with the identification of a standard value and format for the metric in question. The ETL pipeline then transforms the incoming data into a uniform structure. For example, as noted by Interviewee 1, transport companies might report emissions based on various timeframes, such as days, months, or drives. By establishing a standard reporting format (e.g., per month), the data can be transformed in a uniform manner, facilitating aggregation and comparison. However, edge cases, where data for instance spans multiple months, may require more nuanced handling [110]. In extreme situations, data formats may need to be managed manually.

Furthermore, concerns were raised regarding data quality control. First, incorrect data points were mentioned [INT1, INT3], which are difficult to identify and negatively impact the data quality. To mitigate this issue, various strategies can be employed. For internal data, an automated approach may involve constraint enforcement [110]. These constraints can include value constraints, restricting data to a certain range or domain (e.g., ensuring that a price cannot be less than zero) [111]. However, this method is less suitable when dealing with external data, as the data entry process lies outside the control of the core organization. In such a case, a manual approach such as visual analytics can be effective; anomalies or unusual values may be easily identified by a human when represented in charts or diagrams.

Nevertheless, more automated methods, such as entropy analysis or AI methods, can also be employed [110]. Entropy analysis "*measures the level of disorder or uncertainty within a dataset",* by quantifying its unpredictability or impurity [112]. This technique, commonly used in machine learning, applies a mathematical formula that quantifies the dataset's unpredictability. Higher entropy values indicate greater unpredictability [112]. Data points deviating significantly from expected patterns can be flagged as potentially erroneous [113].

As another automated strategy, AI methods can be employed to analyse data in the context of historical trends, thereby flagging data points that appear incorrect based on these patterns [110]. Examples of methods that could be utilized for this purpose are point outlier detection (identifying individual deviating data points) and subsequence outlier detection (identifying sequences of data points that represent an anomaly) [114], [115], [116]. Nevertheless, based on the data specifics certain methods are better applicable than others (e.g., the number of data points for a feature, (in)consistency of frequency, and potential logic behind certain outliers at specific times) [117]. Further research should be dedicated to investigating these models in a CSRD context.

Finally, the issue of varying data hierarchies was identified as a significant challenge for Company E [INT5]. Assuming this challenge is concept-bound (e.g., involving HR, legal entities, or equity shares), Habtezghi suggested that employing a data warehouse could be

an effective solution [110]. In this approach, data is organized according to parent-child relationships, allowing for detailed analysis at various levels. The data is first deconstructed into the smallest coherent units (e.g., an equity share belongs to a legal entity, which in turn belongs to a corporate group). These layers are then consolidated into a warehouse schema, enabling the creation of a unified data hierarchy. This hierarchical structure facilitates the visualization of data at different levels, thus providing comprehensive insights [110].

Phase 5 – Centralizing data

- **Select the appropriate data consolidation technique**: choose between a data warehouse, data lake, or data lakehouse based on the organization's size, data structure needs, and available resources.
- **Address data dispersion**: implement the chosen consolidation tool to centralize data scattered across multiple sources, ensuring a more unified and accessible data environment.
- **Strive for a Single Source of Truth**: integrate external data sources with internal ERP systems and other sources to enhance data quality and reliability, reducing the risks associated with traditional data-sharing methods.
- **Standardize data through ETL processes**: utilize ETL processes to standardize data formats and calculations across internal and external sources. This facilitates easier aggregation, comparison, and analysis.
- **Enhance data quality control**: deploy quality control measures such as automated constraint enforcement, entropy analysis, or AI models to detect and address anomalies in the data.

5.3.6 Phase 6 – Realizing the to-be landscape

Upon the successful connection of all data sources and the centralization of data, Phase 6 of the roadmap focuses on the realization of the to-be landscape. This phase involves integrating the selected tools, such as those for calculating emissions or facilitating reporting, with the central data consolidation platform. For some of these tools, such as the platform by Coolset and Big Mile, data needs to be imported in a certain Excel format. Platforms like a data warehouse or data lake can support a different set of formats; each format needs to be properly specified and implemented. In turn, data can be exported on an Excel sheet with those formatted values. If necessary, rows can be filtered [110]. This solves the issue of manually entering data into an Excel spreadsheet, thereby streamlining the process.

With the to-be landscape now operational, the organization can initiate its integrated reporting processes in line with the CSRD requirements. However, the establishment of this landscape is not the final step. Continuous evaluation and refinement of the IT landscape and underlying structures are crucial to ensuring ongoing compliance, adaptability to regulatory changes, and alignment with organizational goals. Regular assessments should focus on the performance of the integrated tools, data accuracy, and the overall effectiveness of the reporting process.

By maintaining a proactive approach to monitoring and improving the landscape, the organization can ensure that its CSRD reporting remains robust, accurate, and capable of adapting to future demands. In addition, considering the rapid advancements in data management and reporting technologies, it is important for organizations to stay informed about emerging tools and techniques that could further enhance their reporting capabilities.

Phase 6 – Realizing the to-be landscape

- **Integrate selected tools with the centralized platform**: ensure that all chosen tools are fully connected to the central data consolidation platform.
- **Initiate integrated reporting**: once the to-be landscape is operational, begin the process of integrated CSRD reporting, leveraging the newly established data flows and connections.
- **Establish ongoing evaluation processes**: implement regular assessments of the IT landscape, focusing on tool performance, data accuracy, and the effectiveness of the reporting processes to maintain compliance and adaptability.
- **Adopt a proactive monitoring approach**: maintain a proactive stance on monitoring and improving the IT landscape to ensure that CSRD reporting remains robust, accurate, and capable of adapting to future regulatory or operational demands.
- **Stay updated on technological advancements**: monitor emerging tools and techniques in data management and reporting.

5.4 PROPOSED ROADMAP

Based on the aforementioned phases, the final roadmap has been developed and is presented in Figure 20. The steps included in the roadmap are as follows: building support, mapping the IT landscape, creating an overview of the data points, conducting solution research, centralizing data, and realizing the to-be landscape. Additional to the roadmap, a sheet with guidelines has been developed to provide more details to each phase, based on the bold statements in each excerpt. This sheet is included in Appendix 9.10. Together, the roadmap and guidelines zoom in and focus in more detail on "integrated reporting", one of the steps in the roadmap created by BDO (Figure 17).

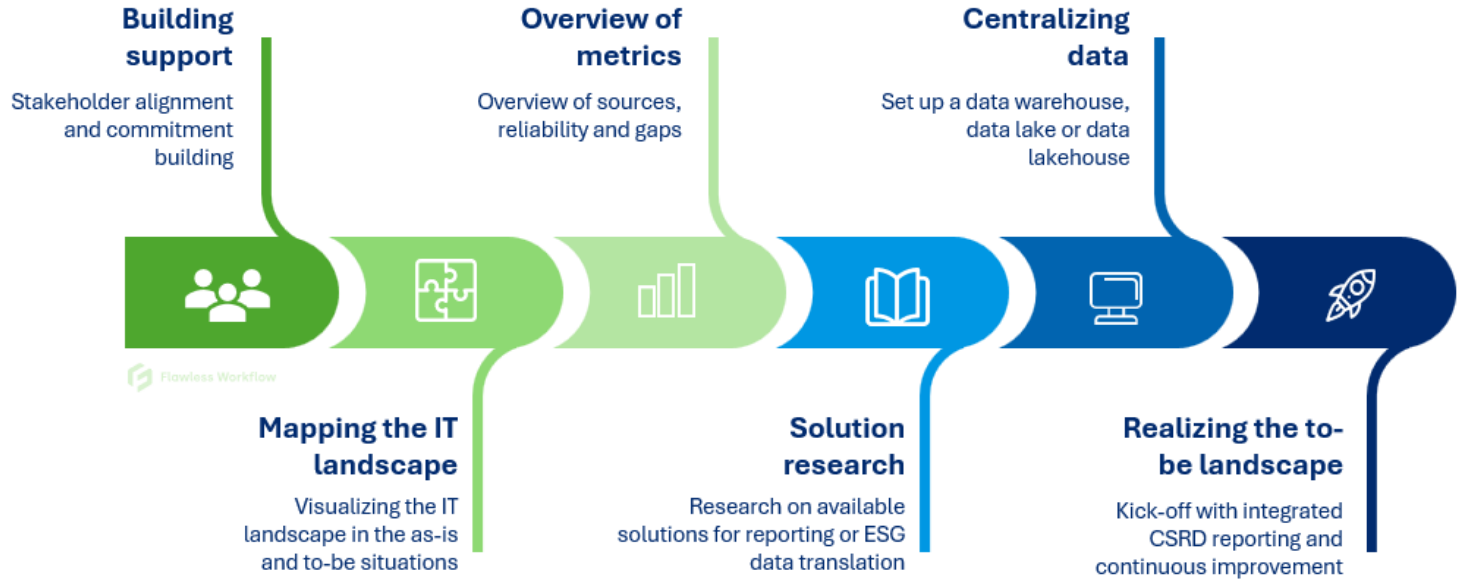


Figure 20 - Roadmap towards integrated CSRD reporting

5.5 DISCUSSION OF ROADMAP

The objective of this study is to address the question:

> *"How can companies integrate data from various sources within their IT architecture and along their value chain, facilitating data consolidation for CSRD reporting?"*.

Following a literature review, we conducted a series of eight interviews with organizations currently engaged in the CSRD implementation process. Based on the challenges and best practices identified in both literature and the interview findings, a roadmap with several guidelines has been developed to address the main research question.

The roadmap starts with building support and commitment for CSRD reporting, as this was highlighted by several interviewees. Internal stakeholders must be informed of CSRD reporting and must be committed to support the CSRD team in various capacities. In essence, this aligns with principles of effective change management, for which additional directions derived from change management literature have been incorporated into the

guidelines. Furthermore, engaging with external stakeholders is vital to facilitate a basis for integrating data from various sources along the value chain. For instance, agreements must be made on data sharing and quality control to ensure the reliability and consistency of the reported data.

Moving to the second phase, the focus lies on mapping the IT landscape. This helps create an understanding of the current state as well as areas where changes are necessary. As the best approach for this task, the literature highlights Enterprise Architecture practices, utilizing frameworks such as TOGAF and modelling languages like ArchiMate. However, we acknowledge that not all organizations possess the required skills or resources to adopt such methodologies. Therefore, the guidelines recognize the use of simpler modelling techniques to achieve similar outcomes. Furthermore, modelling relationships with external stakeholders provides insight into data interactions and how this exchange should be facilitated in a CSRD context. At this stage it is recommended to start thinking of future steps, potentially even beginning with a to-be model.

Phase 3 is based on practical insights from the interviews, which underscored the need to create an overview of metrics derived from the materiality assessment. For each metric, it is valuable to document the data source, the reliability of the source, and availability of data. Additionally, future improvements should be identified and outlined. To supplement this overview, data process flows can be developed to show the "journey" of specific data points throughout the organization's IT landscape and value chain.

Subsequently, the next phase focuses on solution research, which is consistent with the other roadmaps employed by Flawless Workflow. This step involves identifying appropriate tools and applications for CSRD reporting. Insights from the interviews, particularly regarding the software solutions that companies have already adopted, and the challenges encountered during the implementation process, provide valuable input for this phase. Although it may be argued that solution research is not directly related to data consolidation, it is important to have a clear understanding of the tools and applications that will be connected to the data consolidation platform before it is established.

The fifth phase addresses the task of data centralization, for which the literature review showcased a high relevance in the CSRD context. Furthermore, several interviewees expressed using a data consolidation platform for reporting purposes or intent to do so in the future. Based on these findings, we consider that either a data warehouse, data lake, or data lakehouse – depending on the organizational context – can bring great value for CSRD reporting. Such platforms provide the necessary flexibility to adapt to changing regulations, as well as accommodate the interactions with external stakeholders. However, the integration of data from various sources requires careful consideration of the connectivity between these sources and the data consolidation platform. In instances where connectivity presents a challenge, alternative methods such as PDF or Excel exports may be employed, though these approaches may compromise reliability.

As final step, we include the realization of the to-be landscape, where all tools and relevant stakeholders are connected to the data consolidation platform. At this stage, companies can start with integrated CSRD reporting. However, to maintain the quality and effectiveness of the IT landscape and reporting processes, ongoing evaluation and monitoring are essential. Therefore, the roadmap has no specific endpoint as continuous improvement is a critical part of CSRD compliance.

Altogether, these six phases outline how companies can integrate data from various sources within their IT architecture and along their value chain. By following these guidelines, organizations can facilitate data consolidation and enhance their capacity for integrated CSRD reporting. Chapter 6 presents the next stage of the process, which involves the validation of the roadmap using expert opinion and a survey.

6 ROADMAP VALIDATION

Chapter 6 is dedicated to the third step of the Design Cycle: treatment validation. Validation of a treatment, as explained by Wieringa, is "*to justify that it would contribute to stakeholder goals if implemented*" [15, p. 31]. This also incorporates Steps 5 and 6 of the Design Process (Appendix 9.9), which are the presentation of ideas to others to collect feedback and improving the design, respectively [107]. The chosen process to validate the created roadmap is through expert opinion, supported by feedback gathered through a survey. First, in Section 6.1, the procedure followed for the expert opinion is outlined, as well as the findings. Then, the second section of this chapter describes the goals, structure, and outcomes of the survey. The implications for the roadmap are then discussed in Section 6.3, followed by the presentation of the final roadmap and guidelines in Section 6.4.

6.1 EXPERT OPINION

This section describes the validation of the created roadmap via expert opinion. First, we discuss the procedure followed in this process, as well as the questions on which the validation is based. Then, Sub-section 6.1.2 discusses the findings based on these questions.

6.1.1 Procedure

At the core of validation research are a set of questions outlined by Wieringa, divided into four categories: the effects produced by the roadmap, trade-offs (alternative solutions to the roadmap), sensitivity of using the roadmap in different contexts, and requirements satisfaction [15]. These questions are used as a basis for the expert opinion, which is the simplest way to validate an artifact [15]. In this validation model, experts imagine realistic problem contexts and make predictions about the effects of the artifact in said contexts. The set of questions that has been used as guide in the validation process is included in Appendix 9.11.

The roadmap as presented in Figure 20 was submitted to a panel of four experts in consultancy at Flawless Workflow, who individually assessed the effects of the roadmap as well as its guidelines. To ensure a range of responses, experts were chosen with different positions. Their demographics were as follows: BI consultant, CEO (responsible for creating Flawless' other roadmaps), managing consultant, and workflow consultant. Based on their assessment, they answered the questions for each category. During the expert opinion sessions, written notes were taken. Based on these notes, the findings are consolidated in Sub-section 6.1.2.

6.1.2 Expert opinion findings

According to the four categories by Wieringa (effects – requirements – sensitivity – trade-offs) [15], we now discuss the findings of the expert opinion sessions.

Effects

The first category assesses the effects generated by the integration of the roadmap (the artifact) within its intended context. Experts were consulted to evaluate the potential of the roadmap in aiding organizations towards integrated CSRD reporting. Consensus among

the experts indicates that the roadmap provides organizations with a clear and structured approach, thus making the reporting process more manageable. The roadmap's combination of general guidance and specific recommendations for activities or the deployment of methodologies is particularly valuable, as it helps organizations grasp the all-encompassing topic of CSRD. Additionally, the roadmap offers practical guidance at an operational level while simultaneously incorporating strategic management elements. Moreover, the experts highlighted that the roadmap helps consider legal and regulatory requirements throughout the process.

One expert underscored the critical connection between technology and human resources, identifying these as two main factors that must be effectively organized in such a process. Organizations must be aware of their internal structures and processes and rethink these to foster a more integrated approach. Furthermore, it can enhance departmental cohesion, addressing the issue of fragmented practices across different departments. Nevertheless, the experts noted a current gap in the roadmap regarding awareness, training, and comprehension at all organizational levels. To maximize the roadmap's effectiveness, it is essential that individuals understand the broader impact of their actions at every stage of the process.

Based on these points, the experts generally agreed that the roadmap is highly applicable to projects focused on achieving CSRD compliance, regardless of an organization's starting point or specific areas of focus. However, they suggested that the roadmap could be further refined by more explicitly linking certain activities to specific CSRD compliance requirements.

Next, the discussion shifted to what extent the roadmap can be tailored to various industries. The experts acknowledged that in its current form, the roadmap is a solid guideline to provide an industry-wide advice. Moreover, the roadmap is adaptable across different industries, with each sector potentially following a more appropriate route based on the relevance of specific themes (e.g., manufacturing may prioritize different core themes compared to social well-being). Per sector, there are also significant differences between the number of resources they have available, particularly with SMEs generally having fewer resources. One expert proposed the development of a “light” version of the roadmap to accommodate these resource constraints. Nevertheless, even within a single organization, the roadmap is applicable across various scopes.

Shifting the focus inwards of Flawless Workflow, the experts observed that the roadmap holds significant potential for the organization. One expert highlighted that a key target group for Flawless is the social housing sector. Although this sector is not part of the mandatory CSRD compliance group, these organizations have a strong intrinsic motivation to adhere to CSRD principles. Therefore, the roadmap could be beneficial even for organizations motivated by voluntary compliance. Furthermore, the roadmap can serve as a tool for clients to determine their position within the CSRD compliance process, reinforcing the relevance of data maturity assessments.

Finally, the experts were asked whether Flawless Workflow needs to implement organizational changes to effectively embed the roadmap within future projects. They emphasized the importance of enhancing professional knowledge in the areas of CSRD and related software development. The roadmap aligns well with Flawless Workflow's current operational methods, which can facilitate its adoption due to the familiarity of certain practices. The organization already possesses substantial expertise in change management, data centralization, and general reporting.

However, there is a need for more specialized knowledge, particularly in areas such as the double materiality assessment, CSRD metrics, environmental metric calculations, and relevant software and tools. One expert expressed the importance of understanding the overarching vision of CSRD, especially considering aspects that remain undefined. Based on what consultants observe in the field and ongoing development in the CSRD regulation, they should be able to provide clients with informed guidance on their compliance journey.

Requirements

The second category of evaluation considers the roadmap's alignment with the specified requirements. First, the experts were asked whether the roadmap provides clear and actionable guidelines to facilitate the transition towards integrated CSRD reporting. Overall, the experts agreed that the roadmap effectively fulfils its primary purpose, offering steps that are clear, useful, and practical. The guidelines strike an appropriate balance between being "practical" and serving as a "roadmap."

However, the experts noted that for practical application, the roadmap would benefit from incorporating more specific activities or instructions. One expert suggested the inclusion of templates for certain activities to enhance usability. Additionally, complementing the guidelines with sub-steps could offer further clarity on aspects such as ownership and responsibility. Another expert highlighted the need to make the steps more pragmatic and accessible, particularly for client companies that may struggle with certain terminologies.

The second point of evaluation focused on the roadmap's value for consultants working on CSRD compliance projects. The experts highlighted that the roadmap's clear steps facilitate a coherent approach, enabling consultants to guide their clients effectively through the compliance process. It also aids in communicating progress to clients, providing a clear perspective on their current status and future milestones. Internally, the roadmap helps maintain a consistent quality standard. The added value is significant, as consultants accumulate experience within their sector, thereby increasing efficiency in subsequent projects focused on integrated CSRD reporting.

Lastly, the roadmap's value for companies pursuing integrated CSRD reporting independently was considered. The experts agreed that the roadmap offers similar benefits to companies as it does to consultants, primarily in guiding the compliance process. However, it was noted that companies should possess a certain level of in-house technical expertise and IT infrastructure to effectively implement the roadmap's steps. The presence of an innovation manager or a dedicated IT department would likely facilitate the process.

Moreover, developing specific sub-steps and providing clearer guidance on responsibilities could further enhance the roadmap's utility for companies.

Sensitivity

The third category of evaluation concerns the sensitivity of the roadmap, which is its ability to produce desired outcomes across different contexts. The first question asked to the experts was how the roadmap might be adapted for use beyond CSRD compliance. The experts unanimously recognized the potential of the roadmap for general reporting purposes. They agreed that the high-level steps outlined in the roadmap are broadly applicable, though some steps may require minor modifications. For instance, one expert noted that while the overview of metrics is currently tailored to CSRD, particularly with respect to the DMA, these could be easily adapted to define KPIs for general reporting.

The experts then considered the assumptions necessary for the roadmap's application in diverse contexts, such as varying organizational sizes and structures. A key concern was the availability of resources, particularly in terms of time and capacity. Smaller companies, for example, may not need to follow all steps and sub-steps outlined in the roadmap. Additionally, organizations require a certain level of IT maturity and data generation capabilities to effectively implement the roadmap. The ambition of organizations was also highlighted as a critical factor; even with adequate resources, a lack of motivation could hinder the roadmap's implementation.

Given the rapidly evolving regulatory environment of the CSRD, the final question addressed whether the roadmap requires adjustments to maintain its relevance. The primary recommendation was to incorporate a more cyclical approach to the roadmap. Continuous evaluation and monitoring of technological developments should occur concurrently with the implementation of each step, instead of being part only of the final phase. This parallel process would help determine if adjustments are necessary, based on practical experience. Over time, this would allow for the identification of which specific steps are most effective for different types of organizations, enabling the customization of roadmaps tailored to specific organizational structures.

One expert suggested introducing a "Phase 0" that involves setting initial targets based on current laws and regulations. These targets could then be reassessed at the end of the roadmap, with the entire process being repeated if necessary to address any regulatory changes. Another discussed a more agile way of deployment, by incorporating the creation of a proof-of-concept in one of the first sprints, instead of trying to work on everything in one go.

Trade-offs

The fourth category, as outlined by Wieringa [15], considers trade-offs by evaluating whether alternative artifacts in the same context could produce similar effects. The experts were consulted on the existence of other methodologies or tools that could facilitate the transition to integrated CSRD reporting. None of them managed to identify a specific alternative suited to this context. They acknowledged that while there are frameworks

available for designing IT landscapes or establishing reporting architectures through methods like KPI trees, the roadmap developed in this research has a unique position. It distinctively integrates the perspective of digital transformation into the reporting process.

The experts did suggest that other frameworks and tools, such as process mining and the KPI trees, could complement the roadmap as part of a broader solution. All in all, they were unable to offer a view on how the roadmap compares to alternative artifacts. Rather than competing with other tools, the roadmap can be combined with them to enhance the overall process of CSRD reporting.

General remarks

As a concluding point, the experts were invited to share any general remarks. Overall, they expressed that the roadmap has significant potential to support and benefit companies. Particularly, its value would be enhanced if it could be adapted for other reporting purposes. They also noted that the roadmap incorporates elements of Flawless Workflow's standard practices, providing a clear link between their current operations and the CSRD process. The two experts who are part of the management team suggested that the roadmap could be a valuable addition to sales presentations, offering an opportunity to engage with customers and leads about the applicability and perceived value of the guidelines.

6.2 SURVEY

Additional to the validation through expert opinion, feedback was gathered using a survey. This survey was shared with the participants of the interviews, as the roadmap was developed based on their input, and they are familiar with the research context.

6.2.1 Goals and structure

The goal of the survey is to collect feedback on the roadmap and its guidelines, as well as to evaluate their practical applicability. Consistent with the expert opinion validation, the survey is structured according to the categories outlined by Wieringa [15]. The questions initially used for the expert opinion were modified to better align with the needs and experience of the survey participants. Shaping the questions was guided by a "need to know" approach, focusing on essential data required for validating and refining the roadmap and guidelines [118]. This approach was chosen to ensure that the survey remains concise and relevant.

The target audience for the survey comprised the participants from the earlier interviews, who all had expressed a willingness to provide further input during the research process. The survey was composed of Likert scale questions and open-ended text responses [119]. Prior to distribution, the survey questions were reviewed by a supervisor. Based on the received feedback, revisions were made to reduce any ambiguity in the questions. The finalized set of survey questions is included in Appendix 9.12.

The survey sought to achieve several key outcomes. First, to validate the applicability and usefulness of the roadmap and guidelines. Second, to assess whether the roadmap met the criteria of providing clear guidelines and delivering value. Lastly, to determine the

flexibility of the roadmap in a dynamic regulatory environment and its applicability to a broad target group. We anticipated these outcomes to confirm the potential for practical implementation of the roadmap and guidelines.

6.2.2 Conducting the survey

The survey was conducted using Microsoft Forms. Although this method of data collection is typically associated with a lower response rate, it offers the advantages of easy distribution and convenience for participants [118]. The survey was distributed via email, accompanied by an explanation of its purpose and a specified deadline for responses (a range of 8 workdays). Additionally, brief instructions were provided on how to interpret the roadmap (from left to right) and the guidelines (from top to bottom). The conduct of the survey was covered under the ethics approval obtained from the BMS Ethics Committee at the University of Twente.

No alterations were made to the survey questions following their distribution, ensuring that all participants were presented with the same set of questions [120]. The response rate was 62.5%, with 5 out of 8 participants responding, exceeding the minimum acceptable response rate of 40% [121].

6.2.3 Survey outcomes

For the survey, four categories of questions were created. These were the roadmap effectiveness, requirements for the roadmap, flexibility and adaptability, and lastly general remarks. The survey outcomes are discussed according to these categories.

Roadmap effectiveness

The first set of survey questions was designed to assess the effectiveness of the roadmap in guiding companies towards integrated CSRD reporting. The first question explored the perceived usefulness of the roadmap and guidelines. Out of five respondents, three found the roadmap to be “very helpful”, one considered it “somewhat helpful”, and one rated it as “not very helpful”. This respondent later elaborated that their rating was due to a wish for more detailed steps within the process.

Regarding the applicability of the roadmap to CSRD compliance projects, responses were varied: two participants rated it as “highly applicable”, two as “somewhat applicable”, and one remained neutral. When asked about the ease of integrating the roadmap into ongoing projects, one respondent found it “very easy”, two considered it “somewhat easy”, and two expressed a neutral stance. Notably, four out of five respondents indicated that they could effectively utilize the roadmap within their organizations’ ongoing CSRD projects.

Further exploration, with an open question, into necessary adjustments to effectively use the roadmap revealed diverse perspectives. One respondent indicated that the roadmap would have been beneficial to their company, but they found it difficult to assess whether any specific organizational changes were necessary. Another respondent highlighted that their organization sometimes overlooks the importance of a clear overview of metrics, even though it might be an obvious step. A different organization identified the integration of sustainability into existing and new systems as an important change, as such data is

currently dispersed across personal files and various software systems. The final respondent acknowledged the roadmap and guidelines as a *"useful general overview"* during the initial phase of a CSRD compliance project but emphasized the need for more detailed steps for effective implementation.

Lastly, when considering the roadmap's suitability across different industries, three respondents assessed it as fitting "very well", one as "somewhat", and one remained neutral.

Requirements of the roadmap

The next set of questions focused on evaluating whether the proposed roadmap and guidelines meet the predetermined requirements for transitioning to integrated CSRD reporting. The first question examined whether the roadmap and guidelines provide clear steps for this transition. Respondents could choose options ranging from "strongly disagree" to "strongly agree", with three respondents agreeing, one remaining neutral, and one disagreeing.

Following this, the participants were asked how much value the roadmap and guidelines add to their efforts, or those of consultants, in achieving integrated CSRD reporting. One respondent indicated this to be "a lot of value", two stated it adds "some value", one remained neutral, and one indicated it adds "little value".

Two open-ended questions were also posed to the participants. First, we asked participants to identify specific features they considered positive aspects of the roadmap and guidelines. The first respondent highlighted the emphasis on data collection as a significant aspect of the CSRD. The second respondent reiterated the importance of defining metrics and emphasized that "mapping the IT landscape" is a critical task, but not always an obvious step for higher management. The third respondent indicated that *"the roadmap gives a great overview of steps to be taken in the future"* which was echoed by the fourth respondent, who described the roadmap as *"clear, well presented, straightforward"*. The fifth respondent acknowledge that while the roadmap and guidelines offer a high-level general overview, they lack specific instructions on what actions to take and how to execute them.

The second open-ended question asked respondents to identify any features they felt were missing from the roadmap or guidelines. The first respondent noted that the aspects of CSRD that do not involve data are left out. The second respondent suggested that a design phase could be incorporated between centralizing data and realizing the to-be landscape. The fifth respondent referred to their previous comments on the need for more detailed steps, with the third respondent specifically mentioning the need for further elaboration on actions such as "assess potential issues" or "define criteria for solutions". The fourth respondent expressed a desire for a clearer connection to the CSRD requirements, suggesting the inclusion of a step to assess whether all relevant ESRS themes are adequately addressed by the organization. Furthermore, they admitted not being familiar

with the abbreviation "DMA", therefore providing the advice to use the original terms to keep it understandable *"for dummies"* interacting with the roadmap.

Flexibility and adaptability

The next set of questions evaluated the flexibility and adaptability of the roadmap, particularly in the context of the rapidly evolving regulatory environment of the CSRD. When asked whether the roadmap is sufficiently flexible to remain relevant under such conditions, all respondents unanimously answered "Yes".

The respondents were then asked with an open-ended question what conditions are necessary for the roadmap to work in different types of organizations, such as those varying in size or structure. The first respondent wondered about the roadmap's applicability to small enterprises, due to limited resources in terms of full-time employees. The second respondent stated, "*the framework is set up flexible enough to work in different types of organizations"*, while noting that the number of sub-steps may vary depending on the organization's specific context, such as the size of its IT landscape. The third respondent emphasized the need for sufficient resource allocation to sustainability reporting, as the proposed steps would otherwise not be realizable. The fourth respondent complimented the language used in the guidelines, noting that it allows for flexibility, effectively positioning them as true guidelines. Finally, the fifth respondent indicated that the general nature of the roadmap enables it to be applicable across various organizational contexts.

General remarks

As a final question, respondents were invited to share any additional remarks or feedback. The first respondent complimented the roadmap and guidelines, stating that they "*look great"*. The second respondent described the roadmap as an "*overall really good and well-designed framework",* but asked why the solution research is only conducted in the fourth step, instead of for instance immediately after the commitment phase. The third respondent noted that most companies are currently in the initial stage of determining which sustainability data to report. They predicted that in the coming years, as organizations reach a more stable state, the focus will likely shift to improving data quality. To address this evolution, they suggested that additional stages could be incorporated into the roadmap to reflect this progression. The fourth respondent acknowledged the strong content of the roadmap but pointed out that it currently lacks a clear connection to the CSRD requirements, noting that the directive determines the data points and IT ensures it is realized.

6.3 DISCUSSION OF ROADMAP VALIDATION

Both the expert validation and survey outcomes prove the potential for the roadmap and guidelines to be used in practice. The validation shows that the roadmap can be helpful for both consultants and organizations, providing a good guideline towards integrated CSRD reporting. The roadmap is considered suitable across different industries and varying organizational contexts, due to its generalizability. Furthermore, both experts and survey respondents expect the roadmap to remain relevant within the evolving CSRD landscape.

Nevertheless, several areas for improvement were highlighted. First, both the experts and survey participants mentioned that some aspects in the guidelines should be specified more. Considering that it is not possible to include too many details in the scope of a roadmap, only a few steps will be addressed. We acknowledge the notion that specific instructions are needed on what actions to take and how to execute them. Therefore, it will be interesting for future work to develop a more comprehensive package for organizations, including several levels of details to the roadmap and its phases. This can in turn include the additional phases focused on the stable state that organizations will likely reach in the upcoming years. A detail that will be incorporated is the use of original terms instead of abbreviations to keep the guidelines accessible.

One of the experts and one of the survey respondents mentioned that certain CSRD-specific elements can be highlighted more, so that is something that will be included for the revised roadmap and guidelines. Another aspect that was suggested is the inclusion of responsibility. Apart from establishing the cross-departmental reporting team, it is important to determine who is actually responsible for the process towards integrated reporting.

Another suggestion that was given is the development of Phase 0 to incorporate legal and regulatory requirements. Considering that these aspects should be addressed in the general process towards CSRD compliance (e.g., in a general roadmap), this will not be repeated in the roadmap or guidelines. Nevertheless, a statement below the guidelines can attend future users to the fact that some preliminary work is required to commence with this roadmap. This also addresses the statement by one of the survey respondents as to why the aspects of the CSRD that do not involve data are left out.

Additionally, a suggestion was made by one of the survey participants to facilitate a design phase between centralizing the data and realizing the to-be landscape. Considering that such a phase would include an update to the architectural model as well as designing the schemas for the consolidation platform, the decision is made to include these steps within Phase 5 – Centralizing data.

Furthermore, we would like to elaborate on the question why the solution research is conducted in the fourth step, rather than after the commitment phase. Considering the steps in the roadmap, we believe it is important to first establish an understanding of the current IT landscape, including all systems, tools and functionalities. Then, the overview of metrics should first be complete before progressing to the solution research. The rationale behind this order is that specific metrics can dictate the required functionalities in new tools or applications. Therefore, the gaps identified within the existing IT landscape, based on the identified metrics, provide a critical foundation for the solution research. This approach ensures that any new solutions are aligned with the identified needs and gaps, thereby optimizing the selection process.

As part of future practical work, it is an interesting idea to create templates for a more practical set of activities and instructions. This is currently outside of the scope of the roadmap, but they can be based on the details discussed in Section 5.3. Another interesting aspect for future research is the development of more specified roadmaps for certain industries or organization structures. This can help determine what specific route suites a certain organization. Based on practical experience, it can also be possible to work on the suggestion of creating a "light" version of the roadmap for SMEs, addressing the concern by one of the experts and survey respondents that the proposed roadmap would be too extensive for smaller enterprises. Furthermore, future research could explore the application of the proposed roadmap in a more agile manner, examining how it could be structured into sprints. It may be feasible to organize these sprints around specific ESG topics or themes. Additionally, it could improve the adaptability and relevance of the roadmap in its rapidly evolving field.

6.4 FINAL ITERATION OF THE ROADMAP AND GUIDELINES

Considering the insights gathered through the roadmap validation, the decision was made keep the original structure of the roadmap without making modifications. The high-level steps outlined in the roadmap did not require any adjustments, thereby confirming its validity as presented in Figure 20.

However, the accompanying guidelines underwent another iteration of improvement. Several steps were rewritten or newly incorporated to provide greater details. These include an explicit step for creating awareness and understanding at all company levels; addressing responsibilities for the CSRD process; defining the solution research criteria based on gaps in the current IT landscape; and assessing connections and data hierarchies as examples of potential issues. For Phase 5 – Centralizing data, an additional step was introduced to update the to-be model and design the consolidation schema, which may involve the adoption of star or snowflake schema structures.

Furthermore, the revised guidelines underscore the relationship between various phases and steps and the CSRD requirements. For instance, in the context of the overview of metrics, the guidelines now emphasize the necessity of formulating a plan for future data collection as required by the CSRD. Additionally, following the steps related to data centralization, a step has been included to assess the new landscape's compliance with the CSRD requirements. As a final addition to Phase 6 – Realizing the to-be landscape, an assessment has been incorporated on whether the predefined reporting objectives and the required ESRS themes have been adequately covered.

A final significant revision to the guidelines involves the inclusion of a statement clarifying that this roadmap assumes the completion of certain CSRD-related activities beforehand. Furthermore, the roadmap is compatible with broader, more general CSRD roadmaps. The updated guidelines are illustrated in Figure 21.

CSRD ROADMAP – the guidelines

Building support

- Create awareness and understanding at all company levels
- Engage and involve employees; communicate the value of CSRD
- Establish a cross-departmental reporting team and address responsibilities
- Implement effective change management practices
- Standardize data exchange with stakeholders

Mapping the IT landscape

- Visualize the current IT landscape (as-is) to understand data sources
- Use TOGAF & ArchiMate at business, application & technology level
- Incorporate external stakeholders in mapping
- Assess potential issues; e.g., connections and data hierarchies
- Visualize potential future situations (to-be)

Overview of metrics

- Identify required metrics based on the double materiality assessment
- Establish a "Definition of Done"; conduct a gap analysis
- Plan for future data quality improvements
- In case of gaps, plan for future data gathering (required by CSRD)
- Visualize data process flows

Solution research

- Outline gaps between the IT landscape and required functionalities
- Based on previous step, define criteria for solutions
- Explore available tools trough market research; select based on criteria
- Address manual data collection challenges
- Create or update the to-be model; incorporate selected tools

Centralizing data

- Select the appropriate technique (warehouse, lake, lakehouse)
- Update the to-be model and design consolidation schemas
- Address data dispersion and strive for a Single Source of Truth
- Standardize data through ETL & enhance data quality control
- Assess CSRD compliance; the new landscape adheres to requirements

Realizing the to-be landscape

- Integrate selected tools with the centralized platform
- Initiate integrated CSRD reporting and address the impact
- Establish ongoing evaluation processes and proactive monitoring
- Stay updated on technological advancements
- Assess predefined reporting goals & coverage of required ESRS themes

This roadmap assumes that CSRD compliance activities have been performed in advance, such as mapping the value chain, performing a double materiality assessment, and setting goals regarding the legal and regulatory requirements of the CSRD. This roadmap can be used alongside more general CSRD roadmaps.

Figure 21 - Revised roadmap guidelines

7 GENERAL DISCUSSION

This chapter first presents a reflection on aspects of this research in Section 7.1. Then, in Section 7.2, we review the limitations of the study. Section 7.3 is dedicated to exploring avenues of future research. The chapter is finalized with a list of recommendations for Flawless Workflow (Section 7.4).

7.1 CRITICAL REFLECTION

The first point that permits reflection concerns the literature review conducted on the CSRD landscape. During the research process, it became evident that the field is characterized by a high degree of variability and frequent inconsistencies within the literature, likely due to its novel nature. Therefore, we encountered challenges in maintaining an up-to-date and accurate synthesis of the most current developments. Although every effort was made to base the discussion on the most recent advancements available at the time of writing, we acknowledge that the literature review was conducted around March, and subsequent developments may have emerged since then. Consequently, the findings and conclusions presented in this thesis should be viewed within the context of the evolving nature of CSRD research.

Furthermore, selecting the most appropriate data consolidation technique is highly dependent on the specific organizational context, making it challenging to provide a definitive answer to this query. Given the variation in organizational needs and structures, a one-size-fits-all solution is not feasible. Nevertheless, we hope that this study has successfully highlighted key factors that organizations should consider in their decision-making process, thereby offering guidance in selecting the most suitable tool for their circumstances.

Another area that we would like to further evaluate pertains to the questions around "motivation". Motivation is inherently challenging to research and quantity, as individuals may not fully understand or articulate their own motivations. Given this complexity, it might have been more effective to focus on specific objectives or quantifiable goals related to CSRD compliance rather than attempting to study abstract motivational factors. For example, we could have explored the underlying objectives that drive the adoption of CSRD-related software within organizations.

Moreover, one finding that emerged from the interviews was the noticeable hesitance of organizations to adopt software solutions within the CSRD context. This reluctance appears to mirror the ongoing and rapid developments both in the CSRD research field and in the software sector. Additionally, the regulatory environment is continuously evolving, further contributing to this cautious approach. It is our concern that this hesitancy could potentially impact the ability of organizations, particularly smaller ones, to achieve timely CSRD compliance. It is not without reason that several interviewees emphasized the critical importance of starting compliance efforts on time.

Given that the interviews were conducted exclusively with organizations based in the Netherlands, there is a potential for bias in the results that we want to reflect on. The findings may have been shaped by the specific characteristics of Dutch organizational culture, which could differ significantly from the attitudes and practices of organizations in other European countries. These differences may influence various aspects of CSRD implementation, including the approach to digitalization within the CSRD context. As such, the generalizability of the findings to a broader European context may be limited.

An aspect regarding the roadmap that we want to evaluate is the requirement to develop "concrete" guidelines. In retrospect, it may have been ambitious to expect a roadmap to deliver such detailed, actionable guidance, as the nature of a roadmap inherently limits the level of specificity it can provide. To address this challenge, we attempted to mitigate the issue by creating a more extensive version of the roadmap guidelines. Additionally, we have identified and outlined areas for future work, where further efforts should be directed toward investigating specific steps and developing detailed templates. This approach aims to bridge the gap between high-level strategic guidance and the concrete actions required for effective implementation.

Lastly, considering the study field of business and information technology, this thesis contributes to its intersection by addressing the interplay between the managerial aspects of CSRD and the technical requirements of integrated reporting. By bridging this gap, the study provides a foundation for more rounded approaches to CSRD implementation. To the best of our knowledge, this study is the first to consider model-based analysis techniques and data consolidation methods in a CSRD context. Further research, particularly by scholars with a background in computer science, could focus on developing a tool or platform that streamlines the integrated reporting process, based on the roadmap. Such efforts could lead to the creation of a dedicated platform similar to the one created by Coolset, which offers support for CSRD compliance in general.

7.2 LIMITATIONS OF THE STUDY

This study is subject to several limitations that may have influenced its outcomes. Addressing these limitations is important to ensure transparency.

7.2.1 Interview sample

The first limitation of the study pertains to the sample size and composition of the interview participants. The research included only eight organizations, which limits the scope of considered perspectives. Consequently, numerous sectors and types of organizations that are impacted by the CSRD were not represented in the development of the roadmap. To mitigate this limitation, efforts were made to select companies across various industries, including both those mandated to report and those that do so voluntarily. Nevertheless, the relatively small sample may restrict the generalizability of the findings.

7.2.2 Roadmap requirements

Another limitation concerns the source of the roadmap requirements, which were based on general insights from the literature review and empirical interviews. Specific preferences and needs were not directly asked from the interviewees, which may have resulted in the omission of critical aspects. The absence of direct input on specific requirements from participants may limit the roadmap's applicability. Despite this, the roadmap was constructed to comprehensively address the identified challenges, with specific issues being explored in greater depth within Section 1.3. Moreover, insights from the survey were considered for the final version of the roadmap and guidelines to ensure its relevance in practice.

7.2.3 Roadmap development

Furthermore, the absence of a standardized scientific methodology specifically suitable for the creation of roadmaps forms another limitation. This lack of a formalized approach may have resulted in the accidental omission of critical steps, potentially impacting the roadmap's comprehensiveness. To mitigate this risk, steps from the Design Process were systematically incorporated, which is believed to have reduced the potential for such oversights. We do consider the absence of a formalized approach for developing roadmaps to be a gap in research, presenting opportunities for future work. This avenue is further discussed in Section 7.3.

7.2.4 Validation

The final set of limitations relates to the validation process of the roadmap. Due to time constraints, it was not feasible to conduct a second round of interviews with the study participants. Consequently, feedback was gathered through a survey. However, this approach lacked the depth and nuance that direct interviews could have provided, particularly in understanding participants' reasoning behind their responses. Therefore, the absence of detailed qualitative insights may have constrained the depth of the feedback obtained throughout the roadmap validation.

Moreover, the roadmap has not yet been validated through practical application. The chosen validation strategy combined expert opinions with survey feedback. Nevertheless, the application of the roadmap in a real-world project would likely yield more substantive insights into its validity and practical utility. This will therefore be discussed in Section 7.3 on future research.

7.3 FUTURE RESEARCH

There are several avenues to explore in future research, for which ideas emerged during the execution of the study. Accordingly, they will now be presented.

7.3.1 Validating the roadmap

Following validation of the proposed roadmap through expert opinions and a survey, it is important to further validate its practical applicability by implementing the roadmap within the context of real CSRD projects. To ensure a comprehensive validation, future research can for instance investigate applying the roadmap across multiple case studies involving

companies from diverse industries, varying sizes, and distinct organizational structures. This approach will allow for a thorough assessment of the roadmap's effectiveness and adaptability across different environments. By documenting the experiences of companies as they navigate through the roadmap's phases, from initial planning to the final goal of achieving integrated reporting, empirical evidence can be gathered on its utility and impact. This practical validation will not only support the roadmap's credibility but also provide valuable insights into potential refinements needed to enhance its relevance and value in real-world applications.

7.3.2 Agile roadmap

As indicated during the expert validation, a future avenue of exploration is the development of an agile iteration of the proposed roadmap. Such an approach offers organizations the flexibility to implement the roadmap in iterative phases rather than progressing through the entire process in a single, comprehensive effort. This phased approach could mitigate potential challenges associated with attempting to achieve full integrated CSRD reporting in a single instance, thus allowing for a more manageable, adaptive, and responsive adoption of the roadmap.

The potential for an agile framework is particularly relevant given the dynamic and evolving nature of the CSRD landscape as well as organizational environments. Given that there are numerous methodologies for structuring such an agile approach, future research can investigate the most effective frameworks for adapting the roadmap. This creates an understanding of how agile methodologies can be integrated into a structured framework, offering a valuable extension to the proposed roadmap. The next step for this research avenue is to investigate how people can be guided through the roadmap, for example via an interactive tool.

7.3.3 Methodology for roadmap development

Given that organizations such as BDO and Flawless Workflow frequently employ roadmaps to guide their clients, the absence of a scientifically established methodology for constructing such frameworks represents a significant gap in the existing literature. Addressing this gap presents a valuable opportunity for future research to develop a rigorous, evidence-based approach for roadmap creation. This methodology could thereby contribute to both academic knowledge and practical applications. The steps used in this research could serve as a reference point.

7.3.4 Application in alternative contexts

The feedback gathered through the expert opinion indicated the opportunity to modify the roadmap for use in alternative reporting contexts, such as financial reporting. Future research should explore how these adaptations can be effectively implemented, ensuring the roadmap's flexibility and applicability across various reporting scenarios. This investigation can broaden the roadmap’s utility, offering a versatile tool for diverse reporting needs in different organizational settings.

7.3.5 Digital tagging

As none of the interviewees discussed digital tagging, this presents an interesting finding. This omission suggests either that digital tagging is a consideration that may emerge at a later stage in the reporting process, or that it is currently being overlooked. Given that digital tagging is a requirement under the CSRD, further research into this area can be interesting. Future efforts can explore the implications of digital tagging, particularly in relation to the software solutions that support its implementation, the impact on companies, and the potential influence on current CSRD reporting efforts. Such research could provide valuable insights into how organizations can integrate digital tagging effectively, ensuring compliance while optimizing their reporting strategies.

7.3.6 OCR techniques and AI methods

As a final direction for future research, there are several opportunities related to OCR techniques and AI methods. As discussed in Sub-section 5.3.5, current OCR technologies demonstrate limitations in their ability to accurately extract data from tabular formats, such as those found in invoices. This limitation presents an interesting avenue for future research. Investigating how these techniques can be refined to improve accuracy and reliability in such contexts could have great implications for data processing efficiency in various reporting contexts.

Additionally, in Sub-section 5.3.5 we also suggested the potential application of AI methods to identify incorrect data points, thereby eventually enhancing data quality. Future research could focus on identifying and evaluating specific AI models or methods that are most effective in this context. By improving the understanding and application of such technologies, research can contribute to a more robust and reliable data management process.

7.4 RECOMMENDATIONS FOR FLAWLESS WORKFLOW

Based on the comprehensive research conducted in this study, the following recommendations are proposed to Flawless Workflow. First and foremost, it will be a valuable addition for the company to continue to develop and refine the CSRD roadmap, leveraging it as a tool for client projects. Both the literature review and empirical findings highlighted existing gaps in current practices, and the validation of the roadmap demonstrated its potential in a practical context.

However, before fully integrating the roadmap into client projects, Flawless Workflow should prioritize the development of an extensive knowledge base on CSRD-specific issues. This can be achieved by designating one or more employees to specialize in the new directive by delving into its requirements and investigating ongoing developments within the field. By building this expertise, Flawless Workflow will be better prepared to navigate the complexities of CSRD and to provide informed guidance to their clients.

Furthermore, it is recommended that Flawless Workflow begins incorporating the CSRD roadmap into discussions with both existing clients and leads. Engaging in such conversations will help to identify how the roadmap can be effectively embedded in future

projects, thereby facilitating a more practical validation. This process can uncover any missing components or areas for improvement within the roadmap, ensuring that it is comprehensive and tailored to the needs of Flawless Workflow's clients.

Moreover, as highlighted during the expert opinion sessions and in the survey results, the successful implementation of the roadmap in practice requires the development of specific, actionable steps. Future work includes creating a package for companies with such steps. To this end, Flawless Workflow should consider creating standardized templates for each phase and step of the roadmap, customized to align with their existing practices. These templates can serve as practical tools for consultants, ensuring that all necessary actions are systematically addressed throughout the project. The standardization of these steps enhances efficiency and ensures consistency in the application of the roadmap.

The created roadmap is rather generalizable, considering its high-level nature. Nevertheless, as Flawless Workflow begins to use the roadmap across diverse client projects, consultants should document and evaluate specific practices that work well for certain industries or organizational structures. This process of capturing best practices can facilitate creating tailored roadmaps that focus on specific challenges and opportunities for certain industries or structures. The development of these tailored versions can allow Flawless Workflow to deliver more targeted services.

The final recommendation for Flawless Workflow is to actively participate in CSRD-related events, such as conferences, workshops, and seminars. This will enable the company to stay at the forefront of developments within the field, ensuring that consultants stay informed about regulatory changes, best practices, and emerging tools. Furthermore, Flawless Workflow can contribute by sharing insights and experiences while simultaneously learn from the expertise of others.

8 GENERAL CONCLUSION

This final chapter brings the thesis to a close by revisiting the research questions and summarizing the key findings. First, in Section 8.1, we discuss the identified research gap that formed the basis for the research. Following this, Section 8.2 addresses each of the sub-research questions, leading to a discussion of the main research question. Finally, in Section 8.3, we conclude with an overview of the contributions this research makes to both academic knowledge and practice.

8.1 REVISITING THE RESEARCH GAP

This research identified a critical gap in practice and existing literature related to the CSRD, particularly around data integration for CSRD reporting (Section 1.3). With the continuously changing directive and the increasing complexity of reporting requirements, organizations shifted their focus from accountancy towards digitalization of sustainability reporting. This shift reflects a wider challenge in both literature and practice, as organizations struggle to adapt to new reporting methods, make decisions on software and integration, collect data across their value chain, combine qualitative and quantitative data, and ensure software connectivity.

Given these challenges, the literature has so far largely focused on identifying these issues without providing concrete solutions, particularly regarding data integration during the reporting phase. This research addressed this gap by proposing a complete roadmap that offers specific guidelines for data integration. Unlike existing frameworks which mainly guide the general CSRD compliance process, such as the roadmap developed by BDO (Figure 17), this research provides detailed, actionable steps for achieving integrated reporting. This roadmap is intended to help both companies and consultancy firms go through the CSRD implementation process, thereby filling a critical gap in current literature and practice.

8.2 RESEARCH OUTCOMES

To provide an outline of the research outcomes, this section first addresses the sub-research questions posed in Sub-section 1.3.3. Finally, the main research question is answered in Section 8.2.7.

8.2.1 CSRD reporting

Exploring the field of CSRD for RQ 1 has revealed several insights and challenges. The question was as follows:

> *"What does CSRD reporting entail?"*

The literature findings on this topic are presented in Sub-section 3.1.1 and they are discussed in Sub-section 3.2.1. To summarize, the CSRD requires that large organizations, followed by others meeting specific criteria, comply with detailed reporting requirements. These cover environmental, social, and governance topics through 10 topical standards known as the European Sustainability Reporting Standards. These standards require companies to report on their entire value chain, including Scope 3 emissions, with sector-specific standards

expected by mid-2026. Double materiality, including both financial and impact considerations, is central to the CSRD framework.

Moreover, CSRD reporting requires digital tagging and submission in the European Single Electronic Format, using XHTML combined with XBRL tags, and requires external audits to ensure sustainability information matches the quality of financial data. Companies are allowed to hold back sensitive information, provided that the overall relevance of their report is maintained. Relevant KPIs for CSRD reporting are outlined on page 33 of the thesis.

Through the literature review we identified three main categories of challenges that companies face with CSRD reporting. First, there are challenges related to expectations and understanding the CSRD requirements. This demands significant investments of time and resources as well as changes in traditional reporting processes. Secondly, software-related challenges emerge, as traditional methods such as spreadsheets and manual data collection become ineffective, and companies struggle to identify suitable software solutions. Third, data-related challenges include difficulties in reporting across the entire value chain, ensuring data availability and quality, retrieving data from suppliers, and integrating isolated data sources. Additionally, companies face the complex task of combining qualitative and quantitative data, with quantitative data requiring tagging according to digital taxonomy. The most significant challenge identified is software connectivity and integration; ensuring all necessary data and processes are combined within a single reporting system.

8.2.2 Applicable model-based analysis techniques in the context of CSRD

Investigating potential solution areas, we conducted a literature search for RQ 2:

"What model-based analysis techniques can be applied in the context of CSRD?"

The findings (Sub-section 3.1.2) identified Enterprise Architecture as a valuable tool for handle the complexities of CSRD reporting, as discussed in Sub-section 3.2.2. EA provides a helpful framework for visualizing an organization's business and IT landscape, offering stakeholders a complete overview. By leveraging frameworks as implementation guides and using modelling languages for visualizations, EA clarifies the relationships between processes and applications, which is essential for CSRD reporting. This clarity allows organizations to map data relationships and visualize data flows, ensuring a thorough understanding of data origins, application connections, and potential integration pathways. Such a visual overview not only improves the understanding of current operations but also informs strategic decisions, aligning business processes with sustainability goals.

Our findings highlighted that methodologies developed by The Open Group are the most widely recognized model-based analysis techniques for this purpose. In particular, TOGAF and ArchiMate are frequently cited in literature. ArchiMate, noted for its visual expressiveness, is especially relevant in the context of CSRD. Its widespread use in both academic and professional contexts shows its effectiveness in providing a complete view of the CSRD landscape from business, application, and technology perspectives. Therefore,

ArchiMate was selected as the most applicable language for supporting organizations in their CSRD reporting efforts. Tools that support the use of this language are for instance Archi or Bizzdesign.

8.2.3 Data consolidation for CSRD reporting

To address RQ 3:

> *"What are best practices for consolidating data from multiple sources for reporting purposes?",*

Sub-section 3.1.3 explored various data consolidation techniques critical for CSRD reporting, found through an exploratory and systematic literature review. These techniques were discussed in Sub-section 3.2.3. In general, the use of a centralized data source, such as a Single Source of Truth, is essential for ensuring consistency and accuracy throughout the reports, particularly given the extensive data involved in CSRD compliance.

The research identified five key techniques, including ETL (Extract, Transform, Load), ELT (Extract, Load, Transform), data warehouses, data lakes, and the emerging data lakehouse. Considering these techniques in a CSRD context, data warehouses are suitable for smaller organizations with structured data needs, offering user-friendly reporting capabilities. Data lakes are ideal for larger organizations handling both structured and unstructured data, supporting advanced analytics but requiring strong governance. The data lakehouse, combining features of both, offers flexibility and scalability, making it suitable for organizations with the resources and ambition to try new technologies.

Best practices for data consolidation in CSRD reporting depend on the specific needs, ambitions and resources of the organization. Nevertheless, each technique provides a workable option to integrate diverse data sources into a central environment for CSRD reporting purposes.

8.2.4 Motivational drivers for software adaptation and digitalization

Progressing to the practical side of the research, RQ 4 was answered through interviews, as well as its sub-questions 4A and 4B. The interview results are presented in Section 4.3 and the respective questions are discussed in Sub-section 4.4.1.

> *"What are the motivational drivers for companies regarding software adaptation and digitalization in the context of CSRD reporting?"*
>
> - *To what extent are companies motivated to commit to CSRD reporting and what are the drivers behind this motivation?*
> - *How do the choices of companies for strategies and tools reflect their motivation and commitment to digitalization?*

The questions aimed to explore the motivations behind companies' adoption of digitalization for CSRD reporting. The findings reveal a range of motivations among the organizations that participated in the interviews (sub-RQ 4A). Some of the interviewed companies aim to meet only the minimum CSRD requirements, viewing it primarily as a

compliance exercise, while others are motivated to outdo these requirements, driven by a strong commitment to sustainability and the wish to lead in this area. Business value creation, such as cost savings and innovation opportunities, is another significant driver.

Further analysis of companies' digitalization strategies (sub-RQ 4B) generally shows a strong commitment to using digital tools to improve sustainability reporting, whereas with different timelines. While some organizations manage digitalization in-house, others outsource it. Their decisions for tools reflect varying levels of commitment to digitalization. Tools such as Big Mile and Workiva are used for emission calculations and report writing, with several of the interviewed companies combining these tools with data warehouses or data lakes. Nevertheless, some of the participating companies still rely on Excel spreadsheets. Despite varying degrees of adoption, all interviewed companies recognize the importance of transitioning from analogue to digital reporting processes, highlighting the critical role of digitalization in achieving their sustainability goals.

8.2.5 Implementation process for CSRD reporting

RQ 5 addressed the implementation process for CSRD reporting, which is discussed in the interview results (Section 4.3) as well as the discussion of RQ 5 (Sub-section 4.4.2):

> *"What is the implementation process that companies are going through for CSRD reporting?"*

This question examined the diverse approaches organizations take toward CSRD compliance. The process typically begins with the selection of dedicated CSRD teams, which vary in size and departmental representation. Early actions often include creating internal support and conducting materiality assessments, followed by developing technical infrastructure and data management systems.

Challenges are common, with organizations mentioning the complexity, vagueness, and novelty of the CSRD as significant obstacles. Regulatory uncertainties and capacity limitations further complicate for the process, delaying improvements on the digital landscape and creating difficulties in managing extensive stakeholder engagement. Transparency is highlighted as crucial during implementation, alongside the need for incentives to motivate internal stakeholders.

Future milestones of the organizations participating in the interviews focus on completing their materiality assessments and refining digital systems, though they remain cautious due to the changing CSRD. Remarkably, the issue of digital tagging has not yet been addressed, suggesting it may be addressed in later stages or overlooked. Overall, while the implementation process is complex and challenging, progress is being made, with organizations working towards compliance by navigating these obstacles.

8.2.6 Technical constraints to achieve CSRD reporting compliance

Lastly, we explored the technical constraints companies face in achieving CSRD compliance through RQ 6, supported by sub-RQs 6A to 6D.

> *"What are the technical constraints companies face in achieving CSRD reporting compliance?"*
>
> - *What metrics do companies report on in a CSRD context?*
> - *From which sources is the required data obtained?*
> - *To what extent is the necessary data available for companies?*
> - *What are challenges organizations encounter concerning data?*

Throughout the interviews (Section 4.3) the focus was on reported metrics, data sources, availability, and encountered challenges. These were discussed in Sub-section 4.4.3. In summary, companies prioritize certain ESG metrics like emissions, resource use, and workforce conditions.

Data is often sourced from ERP systems like SAP and AFAS, but also involves manual collection from invoices, annual statements, and Excel sheets. Data availability varies widely; while some companies have complete datasets, others only have about half of the required data, often spread across systems.

Key challenges include inconsistent data formats from stakeholders, the need for manual data adjustments, and system limitations not designed for reporting. Data quality is a major concern, with issues like error detection, incomplete data, and reliance on estimations complicating the process. Manual data handling adds to the complexity, being both error-prone and labour-intensive.

Overall, the main technical constraint is integrating fragmented and inconsistent data into a system suitable for CSRD reporting, resulting in significant challenges to compliance efforts.

8.2.7 Integrating data from various sources for CSRD reporting

This section answers the main research question (RQ-O):

> *"How can companies integrate data from various sources within their IT architecture and along their value chain, facilitating data consolidation for CSRD reporting?"*

For this question, the research questions for the literature review (RQ-LR) and empirical chapter (RQ-EC) were used as a basis, which were in turn answered by RQ 1-3 and RQ 4-6, respectively.

- *"What methods and techniques can be utilized to integrate various data sources within companies' IT architecture and along their value chain, facilitating data consolidation for CSRD reporting?"*

- *"What are the practical challenges and experiences that companies encounter while implementing CSRD reporting, particularly in terms of motivational drivers and technical boundaries?"*

The literature review and eight interviews were used to identify challenges and best practices to consider when developing the roadmap (Chapter 5). To answer the main research question, the proposed roadmap provides a structured approach for companies to integrate data from various sources, facilitating integrated CSRD reporting. The roadmap includes six phases (Figure 20), which are described here at a very high level. The validated roadmap and guidelines are presented in Figure 21 in Section 6.4; the entire validation process is discussed in Chapter 6.

- **Building support**: companies begin by creating awareness, understanding and commitment to CSRD reporting with internal and external stakeholders.
- **Mapping the IT landscape**: visualize the current IT landscape using EA practices such as TOGAF and ArchiMate, to understand data sources and potential issues. Investigate a potential to-be landscape.
- **Overview of metrics**: based on the double materiality assessment, create an overview of required metrics, conduct a gap analysis, and plan for future improvements.
- **Solution research**: conduct market research to explore available tools, based on predefined criteria. Address data collection challenges and update the to-be model.
- **Centralizing data**: select a data warehouse, lake or lakehouse and address data dispersion, standardize data, and enhance quality control.
- **Realizing the to-be landscape**: initiate integrated CSRD reporting; establish ongoing evaluation processes and proactive monitoring.

The proposed roadmap offers a comprehensive guide for organizations to streamline their CSRD reporting efforts to integrate data from various sources within their IT architecture and along their value chain. By following these guidelines, organizations can facilitate data consolidation and improve their capabilities for integrated CSRD reporting. This is a unique contribution for both literature and practice, as is elaborated on in Section 8.3.

8.3 CONTRIBUTIONS TO RESEARCH AND PRACTICE

In this final section, we underscore the contributions of this thesis to both research and practice, particularly within the domain of CSRD compliance. Focusing on contributions to research, the study began by identifying a critical gap in the existing literature: while numerous studies have explored the challenge companies encounter in CSRD reporting, few have taken the step to offer concrete, actionable solutions. This thesis directly addresses this gap, offering a novel approach that bridges theoretical understanding with practical implementation.

A key contribution of this research lies in the application and assessment of model-based analysis techniques, alongside data consolidation methods, within the context of CSRD. To the best of our knowledge, this approach represents a novel contribution to the field. By analysing and synthesizing existing models, this study enhances the theoretical framework surrounding CSRD and provides a practical toolkit that organizations can use in their reporting processes.

Moreover, this thesis presents a comprehensive roadmap that guides organizations towards integrated CSRD implementation. This roadmap, informed by the latest academic insights, represents a novel addition to the current body of knowledge. It is designed to be both theoretically strong and practically applicable.

As organizations have progressed further in the CSRD implementation process since earlier research was conducted, this thesis also updates and expands the model of challenges identified in practice. These updated insights reflect the evolving nature of CSRD compliance and offer a more accurate depiction of the current landscape. By doing so, the research not only fills a critical knowledge gap but also lays the foundation for future studies. Several potential areas for further research have been identified in Section 7.3, each of which builds on the novel challenges and findings of in this study.

Considering the practical field, this thesis addresses a critical challenge faced by organizations in the implementation of the CSRD. By analysing the complexities of CSRD, this research contributes valuable knowledge and insights that are essential for navigating the process towards integrated reporting. This thesis begins by offering a comprehensive overview of the CSRD landscape, while additionally discussing relevant model-based analysis techniques and data consolidation methods. These elements form the foundation of the proposed roadmap, which serves as a guide for organizations aiming to overcome the challenges inherent to integrated CSRD reporting.

The proposed roadmap is a central contribution of this thesis, designed to provide actionable guidance for organizations. To ensure its practical relevance and applicability, the roadmap was validated through a dual approach. First, expert opinions were sought. Second, feedback was collected via a targeted survey, enabling the assessment of the roadmap's effectiveness, value, and operational requirements in real-world contexts. This validation process confirmed that the roadmap not only addresses the primary challenges

of CSRD but also offers a framework that can be adopted by organizations across various sectors.

As a final practical contribution, the thesis provides specific recommendations for Flawless Workflow. These recommendations are also broadly applicable and can be leveraged by other consultancy firms engaged in CSRD advisory services. By offering these generalized insights, the thesis extends its utility beyond a single organization, positioning itself as a valuable resource for the wider consultancy industry.

In conclusion, this research makes substantial contributions to both the academic literature and the practical application of CSRD reporting. It offers novel solutions to previously unresolved challenges, enriches theory with updated models, and provides a practical framework that can be directly applied by organizations. The insights and recommendations presented in this thesis can serve as a foundation for both future research and the ongoing practice of CSRD reporting.

REFERENCES

[1] 'Flawless Workflow - Stroomlijn processen, software en data', Flawless Workflow. Accessed: Jul. 10, 2024. [Online]. Available: https://flawlessworkflow.com/

[2] A. Atanasov, 'Digitalization of Sustainability Reporting - Current Trends and Future Problems', in *Cutting-Edge Business Technologies in the Big Data Era*, S. G. Yaseen, Ed., in Studies in Big Data, vol. 135. Cham: Springer Nature Switzerland, 2023, pp. 39–45. doi: 10.1007/978-3-031-42463-2_5.

[3] K. Opferkuch, S. Caeiro, R. Salomone, and T. B. Ramos, 'Circular economy disclosure in corporate sustainability reports: The case of European companies in sustainability rankings', *Sustain. Prod. Consum.*, vol. 32, pp. 436–456, Jul. 2022, doi: 10.1016/j.spc.2022.05.003.

[4] J. Baumüller and S. O. Grbenic, 'Moving from Non-Financial to Sustainability Reporting: Analyzing the EUO Commission's Proposal for a Corporate Sustainability Reporting Directive (CSRD)', *Facta Univ. Ser. Econ. Organ.*, no. 1, p. 369, Dec. 2021, doi: 10.22190/FUEO210817026B.

[5] D. Nooman, 'The Evolution of NFRD to CSRD', Greenomy. Accessed: Feb. 06, 2024. [Online]. Available: https://greenomy.io/blog/evolution-nfrd-csrd

[6] C. de Vries, 'CSRD - Bridging the Gap', Essay (Master), University of Twente, Enschede, The Netherlands, 2023. [Online]. Available: https://essay.utwente.nl/97061/

[7] M. Birkmann, J. Funke, J. Gulbin, L.-M. Meyer, M. Sauer, and L. Wegmann, 'CSRD - Burdening Regulation or Opportunity for CSR Communication? A Qualitative Study on the Influence of the Corporate Sustainability Reporting Directive on Large German Companies', in *Sound or Silence? Current Developments in Organizational Communication*, A. Godulla, M. Ehrlinspiel, S. Gulich, V. Leißner, A. Müller, A. Rüth, and M. Sauer, Eds., Leipzig, 2024, pp. 10–41. [Online]. Available: https://nbn-resolving.org/urn:nbn:de:0168-ssoar-91392-9

[8] A. Aboud, A. Saleh, and Y. Eliwa, 'Does mandating ESG reporting reduce ESG decoupling? Evidence from the European Union's Directive 2014/95', *Bus. Strategy Environ.*, vol. 33, no. 2, pp. 1305–1320, 2023, doi: 10.1002/bse.3543.

[9] S. Pizzi, G. Mastroleo, A. Venturelli, and F. Caputo, 'The digitalization of sustainability reporting processes: A conceptual framework', *Bus. Strategy Environ.*, vol. 33, no. 2, pp. 1040–1050, 2024, doi: 10.1002/bse.3544.

[10] A. A. Vărzaru, 'An Empirical Framework for Assessment of the Effects of Digital Technologies on Sustainability Accounting and Reporting in the European Union', *Electronics*, vol. 11, no. 22, p. 3812, Nov. 2022, doi: 10.3390/electronics11223812.

[11] 'International Study: LeanIX IT Sustainability & ESG Survey Report', LeanIX, 2024. [Online]. Available: https://www.leanix.net/en/download/leanix-it-sustainability-esg-survey-report?utm_term=&utm_source=adwords&utm_medium=ppc&utm_campaign=EMEA-NETHERLANDS_Enterprise-Architecture_AO_PMax_ENG&hsa_ver=3&hsa_cam=21063596320&hsa_grp=&hsa_acc=9751618594&hsa_kw=&hsa_mt=&hsa_net=adwords&hsa_src=x&hsa_tgt=&hsa_ad=&gad_source=1&gclid=CjwKCAjwoJa2BhBPEiwA0l0ImGlTtWvsVs_OJ3EVfoF0W_5G2uCGzT31Eohft2EOuz7eUQNPq8K25hoCr_gQAvD_BwE

[12] P. Svensson, 'Implications of the CSRD on sustainability reporting: the need for digital maturity', School of Industrial Engineering and Management, Stockholm, Sweden, 2023.

Accessed: Feb. 08, 2024. [Online]. Available: https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-329028

[13] A. Spek, N. Sanders, K. Meijer, and W.-J. Dubois, 'The Corporate Sustainability Reporting Directive (CSRD) is already having an impact', PricewaterhouseCoopers (PwC). Accessed: Feb. 26, 2024. [Online]. Available: https://www.pwc.nl/en/topics/sustainability/esg/corporate-sustainability-reporting-directive/the-corporate-sustainability-reporting-directive-csrd-is-already-having-an-impact.html

[14] R. Odobaša and K. Marošević, 'Expected Contributions of the European Corporate Sustainability Reporting Directive (CSRD) to the Sustainable Development of the European Union', presented at the International Scientific Conference "Digitalization and Green Transformation of the EU", 2023. doi: 10.25234/eclic/27463.

[15] R. J. Wieringa, *Design Science Methodology for Information Systems and Software Engineering*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. doi: 10.1007/978-3-662-43839-8.

[16] A. Alturki, G. G. Gable, and W. Bandara, 'The Design Science Research Roadmap: in Progress Evaluation', in *Proceedings of the 17th Pacific Asia Conference on Information Systems (PACIS)*, J. N. Lee, J. Y. Mao, and J. Thong, Eds., 2013, pp. 1–15. [Online]. Available: http://aisel.aisnet.org/

[17] P. Verschuren and H. Doorewaard, *Het ontwerpen van een onderzoek.indd*, 5th ed. Amsterdam: Boom Lemma uitgevers, 2015.

[18] P. Dash, 'Analysis of Literature Review in Cases of Exploratory Research', *SSRN Electron. J.*, 2019, doi: 10.2139/ssrn.3555628.

[19] C. Okoli and K. Schabram, 'A Guide to Conducting a Systematic Literature Review of Information Systems Research', *SSRN Electron. J.*, 2010, doi: 10.2139/ssrn.1954824.

[20] B. Kitchenham and S. Charters, 'Guidelines for performing Systematic Literature Reviews in Software Engineering', Keele University, Keele, EBSE Technical Report, Jan. 2007. [Online]. Available: https://www.researchgate.net/publication/302924724

[21] J. Webster and R. T. Watson, 'Analyzing the Past to Prepare for the Future: Writing a Literature Review', *MIS Q.*, vol. 26, no. 2, pp. xiii–xxiii, 2002.

[22] J. F. Wolfswinkel, E. Furtmueller, and C. P. M. Wilderom, 'Using grounded theory as a method for rigorously reviewing literature', *Eur. J. Inf. Syst.*, vol. 22, no. 1, pp. 45–55, Jan. 2013, doi: 10.1057/ejis.2011.51.

[23] C. Wohlin, 'Guidelines for snowballing in systematic literature studies and a replication in software engineering', in *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, London England United Kingdom: ACM, May 2014, pp. 1–10. doi: 10.1145/2601248.2601268.

[24] V. Garousi, M. Felderer, and M. V. Mäntylä, 'The need for multivocal literature reviews in software engineering: complementing systematic literature reviews with grey literature', in *Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering*, Limerick Ireland: ACM, Jun. 2016, pp. 1–6. doi: 10.1145/2915970.2916008.

[25] J. Lazar, J. H. Feng, and H. Hochheiser, *Research methods in human computer interaction*, 2nd edition. Cambridge, MA: Elsevier, 2017.

[26] V. Braun and V. Clarke, 'Using thematic analysis in psychology', *Qual. Res. Psychol.*, vol. 3, no. 2, pp. 77–101, Jan. 2006, doi: 10.1191/1478088706qp063oa.

[27] M. J. McIntosh and J. M. Morse, 'Situating and Constructing Diversity in Semi-Structured Interviews', *Glob. Qual. Nurs. Res.*, vol. 2, p. 233339361559767, Nov. 2015, doi: 10.1177/2333393615597674.
[28] H. Kallio, A. Pietilä, M. Johnson, and M. Kangasniemi, 'Systematic methodological review: developing a framework for a qualitative semi-structured interview guide', *J. Adv. Nurs.*, vol. 72, no. 12, pp. 2954–2965, Dec. 2016, doi: 10.1111/jan.13031.
[29] 'Corporate sustainability reporting', European Commission. Accessed: Feb. 08, 2024. [Online]. Available: https://finance.ec.europa.eu/capital-markets-union-and-financial-markets/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en
[30] A. Spek and K.-J. de Vries, 'Corporate Sustainability Reporting Directive', PricewaterhouseCoopers (PwC). Accessed: Feb. 06, 2024. [Online]. Available: https://www.pwc.nl/en/topics/sustainability/esg/corporate-sustainability-reporting-directive.html
[31] 'CSRD: wat is het en hoe rapporteert u volgens deze richtlijn', BDO Nederland. Accessed: Feb. 12, 2024. [Online]. Available: https://www.bdo.nl/nl-nl/thema-s/verduurzamen/csrd
[32] 'CSRD en ESRS vragen en antwoorden', Raad voor de Jaarverslaggeving, Nov. 2022. [Online]. Available: https://www.rjnet.nl/siteassets/duurzaamheid/faq-van-rj-en-ser-over-de-csrd-en-esrs.pdf
[33] J.-H. Gnändiger, 'Sustainability reporting: Challenges ahead for corporate leaders', KPMG. Accessed: Feb. 13, 2024. [Online]. Available: https://kpmg.com/xx/en/home/insights/2023/04/ifrs-blog-sustainability-reporting-challenges-corporate-leaders.html
[34] 'Exposure drafts for LSME and VSME', EFRAG. Accessed: Aug. 11, 2024. [Online]. Available: https://efrag-website.azurewebsites.net/lab9?AspxAutoDetectCookieSupport=1
[35] *Commission Delegated Regulation (EU) 2023/2772 of 31 July 2023 supplementing Directive 2013/34/EU of the European Parliament and of the Council as regards sustainability reporting standards*. 2023. Accessed: Feb. 08, 2024. [Online]. Available: http://data.europa.eu/eli/reg_del/2023/2772/oj/eng
[36] '[Draft] EFRAG IG 2: Value chain implementation guidance', EFRAG, Dec. 2023. [Online]. Available: https://efrag.org/Assets/Download?assetUrl=%2Fsites%2Fwebpublishing%2FSiteAssets%2FDraft%2520EFRAG%2520IG%25202%2520VCIG%2520231222.pdf
[37] 'Scope 1 and Scope 2 Inventory Guidance', United States Environmental Protection Agency. Accessed: Feb. 06, 2024. [Online]. Available: https://www.epa.gov/climateleadership/scope-1-and-scope-2-inventory-guidance
[38] P. Bhatia, C. Cummis, A. Brown, D. Rich, L. Draucker, and H. Lahd, *Greenhouse gas protocol: corporate value chain (Scope 3) accounting and reporting standard*. Washington, DC], [Geneva, Switzerland: World Resources Institute ; World Business Council for Sustainable Development, 2011. [Online]. Available: https://ghgprotocol.org/sites/default/files/standards/Corporate-Value-Chain-Accounting-Reporing-Standard_041613_2.pdf
[39] 'The first set of ESRS - the journey from PTF to delegated act (adopted on 31 July 2023) - EFRAG', European Financial Reporting Advisory Group [EFRAG]. Accessed: Feb. 12, 2024. [Online]. Available: https://efrag.org/lab6?AspxAutoDetectCookieSupport=1

[40] 'The European Sustainability Reporting Standards are finalised', PricewaterhouseCoopers [PwC]. Accessed: Feb. 12, 2024. [Online]. Available: https://viewpoint.pwc.com/dt/gx/en/pwc/in_briefs/in_briefs_INT/in_briefs_INT/the-european-sustainability-reporting.html#pwc-topic.dita_84398ad1-a63c-4184-9735-2c0f36b636c2
[41] D. Marvakova, 'CSRD: EU Delays Sector-Specific Sustainability Reporting Standards', cleversoft. Accessed: Feb. 12, 2024. [Online]. Available: https://www.clever-soft.com/25-01-2024-csrd-eu-delays-sector-specific-sustainability-reporting-standards/
[42] 'Commission welcomes agreement on postponing adoption deadlines for certain European Sustainability Reporting Standards', European Commission. Accessed: Feb. 26, 2024. [Online]. Available: https://ec.europa.eu/commission/presscorner/detail/en/mex_24_707
[43] M. Vass, 'Sustainability reporting: MEPs agree with later adoption of standards', News European Parliament. Accessed: Feb. 29, 2024. [Online]. Available: https://www.europarl.europa.eu/news/en/press-room/20240122IPR17036/sustainability-reporting-meps-agree-with-later-adoption-of-standards
[44] '[Draft] EFRAG IG 1: Materiality assessment implementation guidance', EFRAG, Dec. 2023. [Online]. Available: https://efrag.org/Assets/Download?assetUrl=%2Fsites%2Fwebpublishing%2FSiteAssets%2FDraft%2520EFRAG%2520IG%25201%2520MAIG%2520231222.pdf
[45] M. Bossut, I. Jürgens, T. Pioch, F. Schiemann, T. Spandel, and R. Tietmeyer, 'What information is relevant for sustainability reporting? The concept of materiality and the EU Corporate Sustainability Reporting Directive', 2021. [Online]. Available: https://wpsf.de/wp-content/uploads/2021/09/WPSF_PolicyBrief_7-2021_Materiality.pdf
[46] D. Raith, 'The contest for materiality. What counts as CSR?', *J. Appl. Account. Res.*, vol. 24, no. 1, pp. 134–148, 2023, doi: 10.1108/JAAR-04-2022-0093.
[47] K. Staudt, E. Bloem, and M. ter Steege, 'Corporate Sustainability Reporting Directive : An analysis on the pre-implementation phase of CSRD in relation to stakeholder engagement, institutional logic, and networks', Malmö University, Malmö, Sweden, 2023. Accessed: Feb. 13, 2024. [Online]. Available: https://urn.kb.se/resolve?urn=urn:nbn:se:mau:diva-61582
[48] '[Draft] EFRAG IG 3: List of ESRS datapoints – Explanatory note', EFRAG, Dec. 2023. [Online]. Available: https://www.efrag.org/Assets/Download?assetUrl=/sites/webpublishing/SiteAssets/Draft%2520EFRAG%2520IG%25203%2520DPs%2520explanatory%2520note%2520231222.pdf
[49] M. Bolkenstein and L. Lustermans, 'Tentative agreement on the European Single Access Point (ESAP) | Meijburg & Co Tax & Legal', KPMG Meijburg & Co. Accessed: Apr. 26, 2024. [Online]. Available: https://meijburg.com/news/tentative-agreement-european-single-access-point-esap
[50] 'Draft EFRAG IG 3 List of ESRS Data Points'. Oct. 25, 2023. [Excel]. Available: https://www.efrag.org/Meetings/2302241032237237/EFRAG-SRB-Meeting-25-October

[51] T. Passalari, E. Hart, and P. Alexandratou, 'CSRD: The challenges for companies and how to overcome them', dss+. Accessed: Feb. 13, 2024. [Online]. Available: https://www.consultdss.com/content-hub/csrd-challenges-for-companies-and-how-to-overcome-them/

[52] P. Mattila and S. Sasi, 'Towards a Sustainable Future: An Examination of Corporate Responses to the CSRD in the Real-Estate Sector', Master thesis, Uppsala University, Uppsala, Sweden, 2023. [Online]. Available: https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1833419&dswid=-5437

[53] D. Greiling and P. Bauer, 'Sustainability Reporting of State-Owned Enterprises', in *The Routledge handbook of green finance*, 1st ed., O. M. Lehner, T. Harrer, H. Silvola, and O. Weber, Eds., in Routledge international handbooks. , London: Routledge, 2024, pp. 431–445.

[54] M. Eklund and J. A. Vaaler, 'The Transition to CSRD and ESRS', Master thesis, University of Agder, Kristiansand, Norway, 2023. Accessed: Feb. 13, 2024. [Online]. Available: https://uia.brage.unit.no/uia-xmlui/handle/11250/3079855

[55] K. Markendahl, 'An Assessment of Swedish Construction Companies Preparedness for the CSRD', KTH Royal Institute of Technology, Stockholm, Sweden, 2023. Accessed: Feb. 13, 2024. [Online]. Available: https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-330795

[56] M. Tesselaar and L. Huis in het Veld, 'The CSRD requires better ESG data and supporting IT', PricewaterhouseCoopers [PwC]. Accessed: Feb. 13, 2024. [Online]. Available: https://www.pwc.nl/en/topics/sustainability/esg/corporate-sustainability-reporting-directive/csrd-data-tech.html

[57] 'Corporate Sustainability Reporting: data challenges', eurofi, Paris, France, Feb. 2022. [Online]. Available: https://www.eurofi.net/wp-content/uploads/2022/05/the-eurofi-high-level-seminar_paris_corporate-sustainability-reporting_data-challenges_summary_february-2022.pdf

[58] P. Bauer and D. Greiling, 'Greening Austrian social service and healthcare non-profits', *HELIYON*, vol. 10, no. 1, p. e23767, Jan. 2024, doi: 10.1016/j.heliyon.2023.e23767.

[59] N. Glaveli, M. Alexiou, A. Maragos, A. Daskalopoulou, and V. Voulgari, 'Assessing the Maturity of Sustainable Business Model and Strategy Reporting under the CSRD Shadow', *J. Risk Financ. Manag.*, vol. 16, no. 10, 2023, doi: 10.3390/jrfm16100445.

[60] A. Q. Gill, 'Agile enterprise architecture modelling: Evaluating the applicability and integration of six modelling standards', *Inf. Softw. Technol.*, vol. 67, pp. 196–206, Nov. 2015, doi: 10.1016/j.infsof.2015.07.002.

[61] M. M. Lankhorst, 'Enterprise architecture modelling—the issue of integration', *Adv. Eng. Inform.*, vol. 18, no. 4, pp. 205–216, Oct. 2004, doi: 10.1016/j.aei.2005.01.005.

[62] F. Gampfer, A. Jürgens, M. Müller, and R. Buchkremer, 'Past, current and future trends in enterprise architecture—A view beyond the horizon', *Comput. Ind.*, vol. 100, pp. 70–84, Sep. 2018, doi: 10.1016/j.compind.2018.03.006.

[63] M. Välja, R. Lagerström, M. Ekstedt, and M. Korman, 'A Requirements Based Approach for Automating Enterprise IT Architecture Modeling Using Multiple Data Sources', in *2015 IEEE 19th International Enterprise Distributed Object Computing Workshop*, Sep. 2015, pp. 79–87. doi: 10.1109/EDOCW.2015.33.

[64] P. Bhattacharya, 'Modelling Strategic Alignment of Business and IT through Enterprise Architecture: Augmenting Archimate with BMM', *Procedia Comput. Sci.*, vol. 121, pp. 80–88, Jan. 2017, doi: 10.1016/j.procs.2017.11.012.

[65] C. Richter and T. Schaaf, 'A maturity model for tool landscapes of IT service providers', in *12th IFIP/IEEE International Symposium on Integrated Network Management (IM 2011) and Workshops*, May 2011, pp. 1050–1057. doi: 10.1109/INM.2011.5990503.
[66] Z. Zhou, Q. Zhi, S. Morisaki, and S. Yamamoto, 'A Systematic Literature Review on Enterprise Architecture Visualization Methodologies', *IEEE Access*, vol. 8, pp. 96404–96427, 2020, doi: 10.1109/ACCESS.2020.2995850.
[67] I. Abunadi, 'Enterprise Architecture Best Practices in Large Corporations', *Information*, vol. 10, no. 10, Art. no. 10, Oct. 2019, doi: 10.3390/info10100293.
[68] L. Dobrica and E. Niemela, 'A survey on software architecture analysis methods', *IEEE Trans. Softw. Eng.*, vol. 28, no. 7, pp. 638–653, Jul. 2002, doi: 10.1109/TSE.2002.1019479.
[69] R. Capilla, A. Jansen, A. Tang, P. Avgeriou, and M. A. Babar, '10 years of software architecture knowledge management: Practice and future', *J. Syst. Softw.*, vol. 116, pp. 191–205, Jun. 2016, doi: 10.1016/j.jss.2015.08.054.
[70] M. Rohloff, 'An integrated view on business- and IT-architecture', in *Proceedings of the 2008 ACM symposium on Applied computing*, in SAC '08. New York, NY, USA: Association for Computing Machinery, Mar. 2008, pp. 561–565. doi: 10.1145/1363686.1363822.
[71] H. P. Breivold, I. Crnkovic, and M. Larsson, 'A systematic review of software architecture evolution research', *Inf. Softw. Technol.*, vol. 54, no. 1, pp. 16–40, Jan. 2012, doi: 10.1016/j.infsof.2011.06.002.
[72] K. Hinkelmann and A. Pasquini, 'Supporting Business and IT Alignment by Modeling Business and IT Strategy and Its Relations to Enterprise Architecture', in *2014 Enterprise Systems Conference*, Aug. 2014, pp. 149–154. doi: 10.1109/ES.2014.65.
[73] M. J. Anwar and A. Q. Gill, 'A Review of the Seven Modelling Approaches for Digital Ecosystem Architecture', in *2019 IEEE 21st Conference on Business Informatics (CBI)*, Jul. 2019, pp. 94–103. doi: 10.1109/CBI.2019.00018.
[74] The Open Group, 'TOGAF® Standard — Introduction', The TOGAF® Standard. Accessed: Mar. 27, 2024. [Online]. Available: https://pubs.opengroup.org/togaf-standard/adm/chap01.html
[75] B. Pittl and D. Bork, 'Modeling digital enterprise ecosystems with archimate: A mobility provision case study', *Lect. Notes Comput. Sci. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinforma.*, vol. 10371 LNCS, pp. 178–189, 2017, doi: 10.1007/978-3-319-61240-9_17.
[76] M. Kaczmarek-Heß and S. de Kinderen, 'A Multilevel Model of IT Platforms for the Needs of Enterprise IT Landscape Analyses', *Bus. Inf. Syst. Eng.*, vol. 59, no. 5, pp. 315–329, 2017, doi: 10.1007/s12599-017-0482-4.
[77] 'ArchiMate® 2.1 Specification', The Open Group. Accessed: Mar. 13, 2024. [Online]. Available: https://pubs.opengroup.org/architecture/archimate2-doc/
[78] 'About the Service Oriented Architecture Modeling Language Specification Version 1.0.1'. Accessed: Mar. 13, 2024. [Online]. Available: https://www.omg.org/spec/SoaML/
[79] P. Powell and I. Smalley, 'What is data consolidation?', IBM. Accessed: Feb. 15, 2024. [Online]. Available: https://www.ibm.com/topics/data-consolidation
[80] K. Ahmad and S. Z. S. Itmin, 'The Success Factors of Extract-Load-Transform Process in Data Integration Implementation', *J. Inf. Syst. Technol. Manag.*, vol. 7, no. 27, pp. 243–256, Sep. 2022, doi: 10.35631/JISTM.727019.
[81] J. Sreemathy, R. Brindha, M. Selva Nagalakshmi, N. Suvekha, N. Karthick Ragul, and M. Praveennandha, 'Overview of ETL Tools and Talend-Data Integration', in *2021 7th*

International Conference on Advanced Computing and Communication Systems (ICACCS), Mar. 2021, pp. 1650–1654. doi: 10.1109/ICACCS51430.2021.9441984.

[82] E. M. Haryono, Fahmi, A. S. Tri W, I. Gunawan, A. Nizar Hidayanto, and U. Rahardja, 'Comparison of the E-LT vs ETL Method in Data Warehouse Implementation: A Qualitative Study', in *2020 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, Nov. 2020, pp. 115–120. doi: 10.1109/ICIMCIS51567.2020.9354284.

[83] M. Muniswamaiah, T. Agerwala, and C. Tappert, 'Data Virtualization for Analytics and Business Intelligence in Big Data', in *9th International Conference on Computer Science, Engineering and Applications (CCSEA 2019)*, Aircc Publishing Corporation, Jul. 2019, pp. 297–302. doi: 10.5121/csit.2019.90925.

[84] M. Patel and D. B. Patel, 'Progressive Growth of ETL Tools: A Literature Review of Past to Equip Future', in *Rising Threats in Expert Applications and Solutions*, V. S. Rathore, N. Dey, V. Piuri, R. Babo, Z. Polkowski, and J. M. R. S. Tavares, Eds., in Advances in Intelligent Systems and Computing. Singapore: Springer, 2021, pp. 389–398. doi: 10.1007/978-981-15-6014-9_45.

[85] J. C. Nwokeji and R. Matovu, 'A Systematic Literature Review on Big Data Extraction, Transformation and Loading (ETL)', in *Intelligent Computing*, K. Arai, Ed., in Lecture Notes in Networks and Systems. Cham: Springer International Publishing, 2021, pp. 308–324. doi: 10.1007/978-3-030-80126-7_24.

[86] S.-S. Guo, Z.-M. Yuan, A.-B. Sun, and Q. Yue, 'A New ETL Approach Based on Data Virtualization', *J. Comput. Sci. Technol.*, vol. 30, no. 2, pp. 311–323, Mar. 2015, doi: 10.1007/s11390-015-1524-3.

[87] A. Nambiar and D. Mundra, 'An Overview of Data Warehouse and Data Lake in Modern Enterprise Data Management', *Big Data Cogn. Comput.*, vol. 6, no. 4, Art. no. 4, Dec. 2022, doi: 10.3390/bdcc6040132.

[88] A. A. Harby and F. Zulkernine, 'From Data Warehouse to Lakehouse: A Comparative Review', in *2022 IEEE International Conference on Big Data (Big Data)*, Dec. 2022, pp. 389–395. doi: 10.1109/BigData55660.2022.10020719.

[89] K. U. Rehman, U. Ahmad, and S. Mahmood, 'A Comparative Analysis of Traditional and Cloud Data Warehouse', *VAWKUM Trans. Comput. Sci.*, vol. 15, p. 34, Mar. 2018, doi: 10.21015/vtcs.v15i1.487.

[90] M. E. M. El Aissi *et al.*, 'Data Lake Versus Data Warehouse Architecture: A Comparative Study', in *WITS 2020*, S. Bennani, Y. Lakhrissi, G. Khaissidi, A. Mansouri, and Y. Khamlichi, Eds., in Lecture Notes in Electrical Engineering. Singapore: Springer, 2022, pp. 201–210. doi: 10.1007/978-981-33-6893-4_19.

[91] P. Wieder and H. Nolte, 'Toward data lakes as central building blocks for data management and analysis', *Front. Big Data*, vol. 5, p. 945720, Aug. 2022, doi: 10.3389/fdata.2022.945720.

[92] J. Kutay, 'Data Warehouse vs. Data Lake vs. Data Lakehouse: An Overview of Three Cloud Data Storage Patterns', Striim. Accessed: Mar. 20, 2024. [Online]. Available: https://www.striim.com/blog/data-warehouse-vs-data-lake-vs-data-lakehouse-an-overview/

[93] N. E. Janssen, 'The Evolution of Data Storage Architectures : Examining the Value of the Data Lakehouse', Essay (Master), University of Twente, Enschede, The Netherlands, 2022. Accessed: Feb. 15, 2024. [Online]. Available: https://essay.utwente.nl/92801/

[94] B. Shiyal, 'Modern Data Warehouses and Data Lakehouses', in *Beginning Azure Synapse Analytics: Transition from Data Warehouse to Data Lakehouse*, Mumbai, India: Springer, 2021, pp. 21–48. doi: 10.1007/978-1-4842-7061-5.

[95] C. Giebler, C. Gröger, E. Hoos, H. Schwarz, and B. Mitschang, 'Leveraging the Data Lake: Current State and Challenges', C. Ordonez, I.-Y. Song, G. Anderst-Kotsis, A. M. Tjoa, and I. Khalil, Eds., in Lecture Notes in Computer Science, vol. 11708. Cham: Springer International Publishing, 2019, pp. 179–188. doi: 10.1007/978-3-030-27520-4_13.

[96] F. Zouari, N. Kabachi, K. Boukadi, and C. G. Guegan, 'Data Management in the Data Lake: A Systematic Mapping', in *IDEAS 2021: 25TH INTERNATIONAL DATABASE ENGINEERING & APPLICATIONS SYMPOSIUM*, B. C. Desai, Ed., in International Database Engineering and Applications Symposium - Proceedings. New York: Assoc Computing Machinery, 2021, pp. 280–284. doi: 10.1145/3472163.3472173.

[97] T. S. Hukkeri, V. Kanoria, and J. Shetty, 'A Study of Enterprise Data Lake Solutions', *Int. Res. J. Eng. Technol.*, vol. 7, no. 5, pp. 1924–1929, May 2020.

[98] B. Stein and A. Morrison, 'The enterprise data lake: Better integration and deeper analytics', *Technology Forecast: Rethinking integration (PwC)*, no. 1, p. 9, Aug. 20, 2014.

[99] R. Hai, C. Koutras, C. Quix, and M. Jarke, 'Data Lakes: A Survey of Functions and Systems', *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 12, pp. 12571–12590, Dec. 2023, doi: 10.1109/TKDE.2023.3270101.

[100] Z. Chen, 'Observations and Expectations on Recent Developments of Data Lakes', *Procedia Comput. Sci.*, vol. 214, pp. 405–411, Jan. 2022, doi: 10.1016/j.procs.2022.11.192.

[101] D. Oreščanin and T. Hlupić, 'Data Lakehouse - a Novel Step in Analytics Architecture', in *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, Sep. 2021, pp. 1242–1246. doi: 10.23919/MIPRO52101.2021.9597091.

[102] B. Lorica, M. Armbrust, R. Xin, M. Zaharia, and A. Ghodsi, 'What Is a Lakehouse?', Databricks. Accessed: Mar. 20, 2024. [Online]. Available: https://www.databricks.com/blog/2020/01/30/what-is-a-data-lakehouse.html

[103] D. Mazumdar, J. Hughes, and J. B. Onofre, 'The Data Lakehouse: Data Warehousing and More', Oct. 12, 2023, *arXiv*: arXiv:2310.08697. doi: 10.48550/arXiv.2310.08697.

[104] M. Zaharia, A. Ghodsi, R. Xin, and M. Armbrust, 'Lakehouse: A New Generation of Open Platforms that Unify Data Warehousing and Advanced Analytics', presented at the Conference on Innovative Data Systems Research, 2021. Accessed: Mar. 20, 2024. [Online]. Available: https://www.semanticscholar.org/paper/Lakehouse%3A-A-New-Generation-of-Open-Platforms-that-Zaharia-Ghodsi/451cf5fc9786ed4f7e1d9877f08d00f8b1262121

[105] C. McMullin, 'Transcription and Qualitative Methods: Implications for Third Sector Research', *Volunt. Int. J. Volunt. Nonprofit Organ.*, vol. 34, no. 1, pp. 140–153, Feb. 2023, doi: 10.1007/s11266-021-00400-3.

[106] A. Kraai and E. Coenders, 'Aan de slag met de CSRD', Enschede, The Netherlands, Jun. 06, 2024.

[107] T. Munro, 'Appreciative Co-design: From Problem Solving to Strength-Based Re-authoring in Social Design', presented at the Design Research Society, Mar. 2016. doi: 10.21606/drs.2016.271.

[108] J. S. Oakland and S. Tanner, 'Successful Change Management', *Total Qual. Manag. Bus. Excell.*, vol. 18, no. 1–2, pp. 1–19, Mar. 2007, doi: 10.1080/14783360601042890.

[109] L. L. De Oliveira *et al.*, 'Evaluating and mitigating the impact of OCR errors on information retrieval', *Int. J. Digit. Libr.*, vol. 24, no. 1, pp. 45–62, Mar. 2023, doi: 10.1007/s00799-023-00345-6.
[110] X. Fransisco and A. Habtezghi, 'Technical data session', Aug. 14, 2024.
[111] D. Weber, 'A Constraint-Based Approach to Data Quality in Information Systems', Thesis, University of Zurich, Zurich, 2017. [Online]. Available: https://www.research-collection.ethz.ch/bitstream/handle/20.500.11850/198644/1/Thesis-DavidWeber.pdf
[112] 'Entropy in Machine Learning', Deepgram. Accessed: Aug. 16, 2024. [Online]. Available: https://deepgram.com/ai-glossary/entropy
[113] A. Fou, 'Entropy Analysis—How We Use It For Bot (And Human) Detection', Forbes. Accessed: Aug. 16, 2024. [Online]. Available: https://www.forbes.com/sites/augustinefou/2020/09/20/entropy-analysishow-we-use-it-for-bot-detection/
[114] A. Blázquez-García, A. Conde, U. Mori, and J. A. Lozano, 'A Review on Outlier/Anomaly Detection in Time Series Data', *ACM Comput Surv*, vol. 54, no. 3, p. 56:1-56:33, Apr. 2021, doi: 10.1145/3444690.
[115] M. Gupta, J. Gao, C. C. Aggarwal, and J. Han, 'Outlier Detection for Temporal Data: A Survey', *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 9, pp. 2250–2267, Sep. 2014, doi: 10.1109/TKDE.2013.184.
[116] K. Choi, J. Yi, C. Park, and S. Yoon, 'Deep Learning for Anomaly Detection in Time-Series Data: Review, Analysis, and Guidelines', *IEEE Access*, vol. 9, pp. 120043–120065, 2021, doi: 10.1109/ACCESS.2021.3107975.
[117] F. Mesquita, 'AI models in CSRD reporting context (personal communication)', Aug. 28, 2024.
[118] T. Jones, M. Baxter, and V. Khanduja, 'A quick guide to survey research', *Ann. R. Coll. Surg. Engl.*, vol. 95, no. 1, pp. 5–7, Jan. 2013, doi: 10.1308/003588413X13511609956372.
[119] J. R. Evans and A. Mathur, 'The value of online surveys', *Internet Res.*, vol. 15, no. 2, pp. 195–219, Apr. 2005, doi: 10.1108/10662240510590360.
[120] F. Scheuren, *What is a Survey?* American Statistical Association, 2004. [Online]. Available: https://books.google.nl/books?id=CPqsjgEACAAJ
[121] D. A. Story and A. R. Tait, 'Survey Research', *Anesthesiology*, vol. 130, no. 2, pp. 192–202, Feb. 2019, doi: 10.1097/ALN.0000000000002436.

9 APPENDIX A – SYSTEMATIC LITERATURE REVIEW

9.1 INCLUSION AND EXCLUSION CRITERIA

Table 8 - Inclusion and exclusion criteria

RQ 2	Inclusion	**Language**	English or Dutch
		Date range	2000-2024
		Research area	Information systems, Enterprise Architecture, Computer Science, Software Engineering
		Title-Author-Key	Include the search terms or their derivatives
	Exclusion	**Research area**	Medicine, Social Sciences, Natural Sciences, Mathematics, Environmental Science
		Availability	Not open access
RQ 3	Inclusion	**Language**	English or Dutch
		Date range	2020-2024
		Research area	Information Systems, Enterprise Architecture, Computer Science, Software Engineering, Information Management
		Title-Author-Key	Include the search terms or their derivatives
	Exclusion	**Research area**	Medicine, Social Sciences, Natural Sciences, Mathematics, Environmental Science
		Availability	Not open access or pre-access
		Title-Author-Key	Articles focused on specific data warehouse or data lake applications (e.g., in healthcare)

9.2 SEARCH TERMS

Table 9 - Search terms RQ 2

Approach	**Model-based techniques**	**Domain of interest**
Methodolog*	Model-based analysis	IT landscape
Approach*	Model-based analysis technique*	IT architecture
Method*	Model-based analysis approach*	Application* landscape
Technique*	Model* technique*	IT system* architecture
Framework*	Model* approach*	
	Model* method*	

Table 10 - Search terms RQ 3

Use case	**Context**	**Context**
Use case*	Data lake*	Data warehous*
Business application*	Data lake architecture*	Data warehous* architecture*
Application*	Datalake*	Datawarehous*
Scenario*		

9.3 SEARCH QUERIES

Table 11 - Search Queries RQ 2

IEEExplore	**Document Title**	("Document Title":IT landscape OR "Document Title":IT architecture OR "Document Title":Application* landscape OR "Document Title":IT system* architecture)	22 results 14 papers
		("Document Title":model*)	
	All Metadata	("All Metadata":Model-based analysis OR "All Metadata":Model-based analysis technique* OR "All Metadata":Model-based analysis approach* OR "All Metadata":Model* technique* OR "All Metadata":Model* approach* OR "All Metadata":Model* method*)	
	Publication Topics	Enterprise Architecture, Information Technology, Business Rules, Business Services, Data Sources, Design Science, Design Science Research, IT Systems, Information System, Information Technology Services, Modeling Techniques, Business Processes, Information Technology Infrastructure, Model Architecture	
ACM Digital Library	**Title**	[[Title: it landscape] OR [Title: it architecture] OR [Title: application* landscape] OR [Title: it system* architecture]]	3 results 2 papers
	Abstract	[[Abstract: model*] OR [Abstract: methodolog*] OR [Abstract: approach*] OR [Abstract: method*] OR [Abstract: technique*] OR [Abstract: framework*]]	
	Filters	[E-Publication Date: (01/01/2000 TO 31/12/2024)]	
ISI Web of Science	**Topic**	(TS=("IT landscape" OR "IT architecture" OR "application* landscape" OR "it system* architecture"))	13 results 6 papers
		TS=("model-based analysis" OR "model-based technique*" OR "model-based approach*" OR "model* technique*" OR "model* approach*" OR "model* method*"))	
		TS=(methodolog* OR approach* OR method* OR technique* OR framework*)	
Google Scholar	**Title**	(Methodolog* OR Approach* OR Method* OR Technique* OR Framework*)	20.000 results 7 papers
		(Model-based analysis OR Model-based analysis technique* OR Model-based analysis approach* OR Model* technique* OR Model* approach* OR Model* method*)	
		(IT OR IT landscape OR IT architecture OR Application* landscape OR IT system* architecture)	
ScienceDirect	**Full text**	(methodology OR approach OR method OR technique OR framework) NOT machine learning	20 results 7 papers
	Title – Abstract – Keywords	("IT landscape" OR "IT architecture" OR "application landscape" OR "it system architecture") AND ("model-based analysis" OR "model-based technique" OR "model-based approach" OR "modelling technique" OR "modelling approach")	
	Title	("IT landscape" OR "IT architecture" OR "application landscape" OR "it system architecture")	
	Subject area	Computer Science	
	Publication title	Procedia Computer Science, Journal of Systems and Software, Information and Software Technology, Data & Knowledge Engineering, Journal of Systems Architecture, Design Studies (NOT graphics, AI, medical, environmental, etc.)	
Scopus	**Title – Abstract – Keywords**	"IT landscape" OR "IT architecture" OR "application* landscape" OR "it system* architecture"	25 results 10 papers
		"model-based analysis" OR "model-based technique*" OR "model-based approach*" OR "model* technique*" OR "model* approach*" OR "model* method*"	
		methodolog* OR approach* OR method* OR technique* OR framework*	

Table 12 - Search Queries RQ 3

IEEExplore	**All Metadata**	"Use case*" OR "Business application*" OR Application* OR Scenario*	23 results 8 papers
		"Data lake*" OR "Data lake architecture*" OR Datalake*	
		"Data warehous*" OR "Data warehous* architecture*" OR Datawarehous*	
	Document Title	"Data lake*" OR Datalake* OR "Data warehous*" OR Datawarehous*	
ACM Digital Library	**Title**	[[Title: "data lake*"] OR [Title: datalake*] OR [Title: "data warehous*"] OR [Title: datawarehous*]]	9 results 2 papers
	Abstract	[[Abstract: "data lake*"] OR [Abstract: "data lake architecture*"] OR [Abstract: datalake*] OR [Abstract: "data warehous*"] OR [Abstract: "data warehous* architecture*"] OR [Abstract: datawarehous*]]	
	All	[[All: "use case*"] OR [All: "business application*"] OR [All: application*] OR [All: scenario*]]	
	Filters	[E-Publication Date: (01/01/2020 TO 31/12/2024)]	
ISI Web of Science	**Topic**	("Data lake*" OR "Data lake architecture*" or Datalake*) AND ("Use case*" OR "Business application*" OR Application* OR Scenario*) AND ("Data warehous*" OR "Data warehous* architecture*" OR Datawarehous*)	11 results 5 papers
	Title	"Data lake*" OR Datalake* OR "Data warehous*" OR Datawarehous*	
Google Scholar	**Title**	("Use case*" OR "Business application*" OR Application* OR Scenario*)	2.980 results 8 papers
		("Data lake*" OR "Data lake architecture*" or Datalake*)	
		("Data warehous*" OR "Data warehous* architecture*" OR Datawarehous*)	
ScienceDirect	**Full text**	"Use case" OR "Business application" OR Application OR Scenario	35 results 4 papers
	Title – Abstract – Keywords	"Data lake" OR "Data lake architecture" OR Datalake OR "Data Lakehouse" OR "Data warehouse" OR "Data warehouse architecture" OR Datawarehouse	
	Title	"Data lake" OR Datalake OR "Data warehouse" OR Datawarehouse OR "Data Lakehouse"	
	Subject area	Computer Science, Engineering, Management and Accounting	
Scopus	**Title – Abstract – Keywords**	"Use case*" OR "Business application*" OR application* OR scenario*	23 results 7 papers
		"Data lake*" OR "Data lake architecture*" OR datalake*	
		"Data warehous*" OR "Data warehous* architecture*" OR datawarehous*	
	Title	"Data lake*" OR datalake* OR "Data warehous*" OR datawarehous*	

9.4 DISTRIBUTION OF SELECTED LITERATURE OVER YEAR OF PUBLICATION

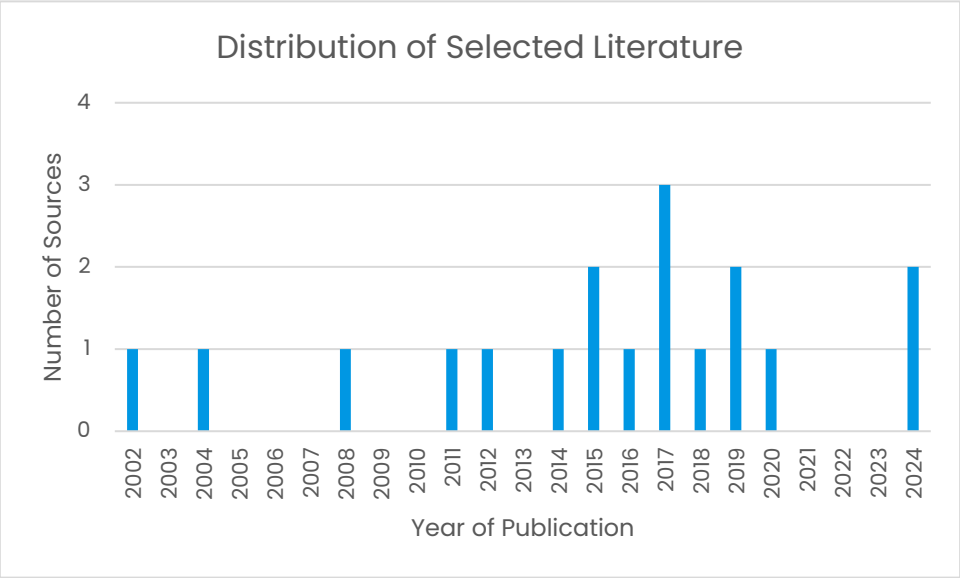


Figure 23 - Distribution of selected literature over year of publication, RQ 2

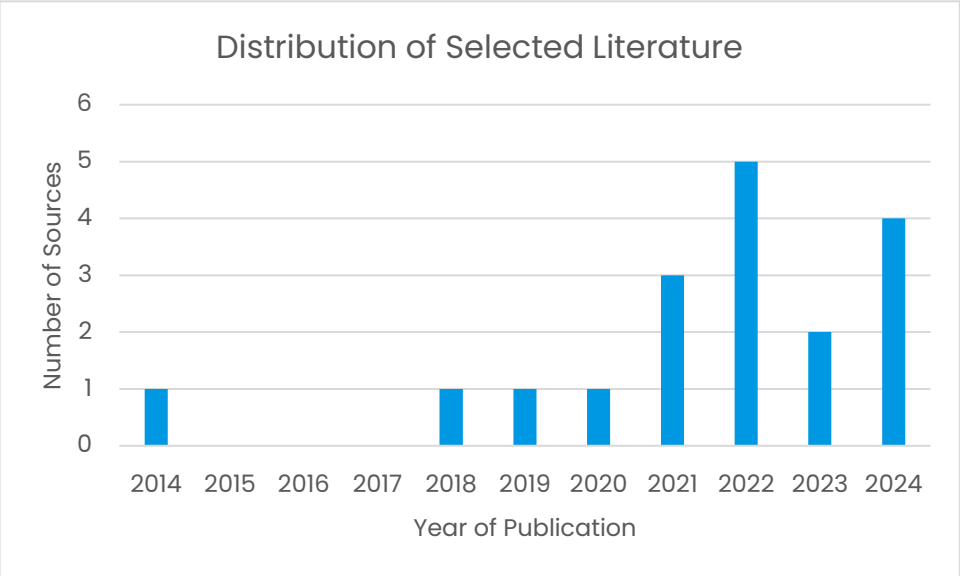


Figure 22 - Distribution of selected literature over year of publication, RQ 3

9.5 CONCEPT MATRICES

Table 13 - Concept Matrix RQ 2

Article	Authors	Concepts				
		Software or Enterprise Architecture	IT Landscape	Architecture Frameworks	Modelling Languages	Modelling Characteristics
An integrated view on business- and IT-architecture	[70] - Rohloff, M.	x	x	x		
A systematic review of software architecture evolution research	[71] - Breivold, H. P., Crnkovic, I., Larsson, M.			x		
A survey on software architecture analysis methods	[68] - Dobrica, L., Niemela, E.	x				
A maturity model for tool landscapes of IT service providers	[65] - Richter, C., Schaaf, T.	x		x	x	
10 years of software architecture knowledge management: Practice and future	[69] - Capilla, R., Jansen, A., Tang, A., Avgeriou, P., Babar, M. A.	x			x	
Agile enterprise architecture modelling: Evaluating the applicability and integration of six modelling standards	[60] - Gill, A. Q.	x		x	x	
Modelling Strategic Alignment of Business and IT through Enterprise Architecture: Augmenting Archimate with BMM	[64] - Bhattacharya, P.	x		x	x	
Supporting Business and IT Alignment by Modeling Business and IT Strategy and Its Relations to Enterprise Architecture	[72] - Hinkelmann, K., Pasquini, A.			x	x	x
A Requirements Based Approach for Automating Enterprise IT Architecture Modeling Using Multiple Data Sources	[63] - Välja, M., Lagerström, R., Ekstedt, M., Korman, M.	x				
A Systematic Literature Review on Enterprise Architecture Visualization Methodologies	[66] - Zhou, Z., Zhi, Q., Morisaki, S., Yamamoto, S.	x			x	
A Multilevel Model of IT Platforms for the Needs of Enterprise IT Landscape Analyses	[76] - Kaczmarek-Heß, M., de Kinderen, S.	x			x	
Modeling digital enterprise ecosystems with archimate: A mobility provision case study	[75] - Pittl, B., Bork, D.			x	x	x
ArchiMate® 2.1 Specification	[77]			x	x	
About the Service Oriented Architecture Modeling Language Specification Version 1.0.1	[78]				x	
Past, current and future trends in enterprise architecture—A view beyond the horizon	[62] - Gampfer, F., Jürgens, A., Müller, M., Buchkremer, R.	x		x	x	x
Enterprise Architecture Best Practices in Large Corporations	[67] - Abunadi, I.	x		x	x	x
A Review of the Seven Modelling Approaches for Digital Ecosystem Architecture	[73] - Anwar, M. J., Gill, A. Q.			x	x	
Enterprise architecture modelling—the issue of integration	[61] - Lankhorst, M. M.	x			x	x

Table 14 - Concept Matrix RQ 3

Articles	Authors	Concepts						
		DW – concept	DW – use cases	ETL & ELT	DL – concept	DL – use cases	DW vs. DL comparison	Data Lakehouse
An Overview of Data Warehouse and Data Lake in Modern Enterprise Data Management	[87] - Nambiar, A. & Mundra, D.	x	x	x	x		x	x
Data Lake Versus Data Warehouse Architecture: A Comparative Study	[90] - El Aissi, M. E. M., Benjelloun, S., Loukili, Y., Lakhrissi, Y., Boushaki, A. E., Chougrad, H., Elhaj Ben Ali, S.	x		x	x		x	
Data Management in the Data Lake: A Systematic Mapping	[96] - Zouari, F., Kabachi, N., Boukadi, K., Guegan, C. G.				x	x	x	
A Study of Enterprise Data Lake Solutions	[97] - Hukkeri, T. S., Kanoria, V., Shetty, J.				x	x	x	
Data Lakes: A Survey of Functions and Systems	[99] - Hai, R., Koutras, C., Quix, C., Jarke, M.			x	x	x		x
Toward data lakes as central building blocks for data management and analysis	[91] - Wieder, P. & Nolte, H.	x			x			x
Modern Data Warehouses and Data Lakehouses	[94] - Shiyal, B.	x		x	x			x
Observations and Expectations on Recent Developments of Data Lakes	[100] - Chen, Z.				x		x	
From Data Warehouse to Lakehouse: A Comparative Review	[88] - Harby, A. A. & Zulkernine, F.	x		x	x		x	x
Data Lakehouse - a Novel Step in Analytics Architecture	[101] - Oreščanin, D. & Hlupić, T.				x			x
The Data Lakehouse: Data Warehousing and More	[103] - Mazumdar, D., Hughes, J., Onofre, J. B.			x				x
The Evolution of Data Storage Architectures : Examining the Value of the Data Lakehouse	[93] - Janssen, N. E.	x			x		x	x
What Is a Lakehouse?	[102] - Lorica, B., Armbrust, M., Xin, R., Zaharia, M., Ghodsi, A.				x		x	x
Data Warehouse vs. Data Lake vs. Data Lakehouse: An Overview of Three Cloud Data Storage Patterns	[92] - Kutay, J.	x			x		x	x
A Comparative Analysis of Traditional and Cloud Data Warehouse	[89] - Rehman, K. U., Ahmad, U., Mahmood, S.	x		x				
The enterprise data lake: Better integration and deeper analytics	[98] - Stein, B. & Morrison, A.				x	x		
Leveraging the Data Lake: Current State and Challenges	[95] - Giebler, C., Gröger, C., Hoos, E., Schwarz, H., Mitschang, B.				x			
Lakehouse: A New Generation of Open Platforms that Unify Data Warehousing and Advanced Analytics	[104] - Zaharia, M., Ghodsi, A., Xin, R., Armbrust, M.						x	x

9.6 INTERVIEW QUESTIONS

Introduction

1. Your position in the company is *...*, can you briefly explain how you are involved with CSRD? *OR* Can you briefly explain your position in the company and how you are involved with CSRD?
2. What is the exact sector under which the organization falls?

Motivation

1. The EU created the CSRD to which your organization must comply. Since when has the organization started working on getting things ready for CSRD compliance?
2. What was the attitude of the organization at the beginning?
    1. And what about individual employees?
3. What is the current level of motivation in general?
4. What is the level of motivation of the organization in terms of digitalization in a CSRD context?
    1. Does the organization consider using software solutions or digital techniques for CSRD reporting?
    2. What kind of solutions have been investigated?
    3. *Follow up if they want to work with Excel:* what do you think about the scalability of this solution?

Implementation Process

1. Since the CSRD was announced, what have the overall steps been to work towards compliance?
2. Have there been efforts to visualize the IT landscape, potentially with a current and to-be situation?
3. Who are the key stakeholders involved in the CSRD reporting process, both internally and externally?
    1. Who are the members of the internal project group?
    2. Are parties in the value chain involved with shaping the process?
4. What specific steps are undertaken to consolidate data for the reports?
    1. Have you encountered any challenges in achieving interoperability? If yes, what?
    2. How have you addressed them?
    3. What tools or platforms are used for data integration and consolidation?
        1. What are positive aspects of these tools?
        2. What are challenges with these tools?
5. What specific IT infrastructure changes or upgrades have been or will be necessary to support CSRD reporting?
6. What are future milestones in your planning related to CSRD?

Technical boundaries

1. What are examples of KPIs or metrics that you are or will be reporting on?
2. What kind of data is necessary to report on these KPIs (data points)?
3. How much of the data points is available for you inside the organization?
    1. What kind of data is this?
    2. From what kind of systems are you retrieving this data?
    3. How is this data exported?
    4. Do you currently experience any issues with retrieving internal data for the CSRD reports?
4. How much of the data points must be retrieved from organizations in your value chain?
    1. What kind of data is this?
    2. From where are you retrieving the data in the value chain?
    3. What are the problems when it comes to retrieving data from external parties?
    4. How are you dealing with these problems?
5. What measures are in place to ensure the quality and accuracy of data used for CSRD reporting?

Lessons learned

1. What lessons have you learned so far from the process of preparing for CSRD compliance?
2. What advice would you give to other organizations embarking on a similar journey?
3. How have you experienced the CSRD implementation process so far in terms of complexity?

9.7 CODING SCHEME

Table 15 presents the coding structure that was developed through a qualitative analysis of the interview transcripts.

Table 15 - Codes per depth level, including their descriptions

Level 1	Level 2	Level 3	Description
IT Landscape	Centralizing the landscape		Based on experiences during the implementation process, companies try to centralize their internal and external IT landscape by making it less complex.
	Consolidation tool	Data lake	A data lake is or will be used as a consolidation tool for reporting purposes.
		Data warehouse	A data warehouse is or will be used as a consolidation tool for reporting purposes.
	Flexibility	Between solutions	Flexibility in the sense of being able to switch between software solutions.
		Within solutions	Flexibility within systems and software solutions.
	Software solutions	Existing systems	Existing software systems that are used in a CSRD context.
		Future expectations	Expectations for future software systems and their evolvement.
	Source systems		Source systems from which data is retrieved for the CSRD reporting.
	Visualizing the landscape		Visualizing the IT landscape, e.g., through an enterprise architecture model or another overview.
CSRD Process	Business value		How sustainability or CSRD compliance can bring business value.
	Central role		The company experiences a coordinating and central position in the CSRD implementation process and cooperation with stakeholders.
	Changes based on insights		Based on insights that are already gathered through the CSRD implementation, changes are made in company conduct.
	Complexity		How and why the CSRD implementation is experienced as complex.
	CSRD timeframe	Actual timeframe	The time that has been spent so far on the implementation process around the CSRD.
		Starting on time (advice)	Lesson learned by companies to really start on time.
	Double materiality	Own assessment	The company is conducting their own internal assessment for the double materiality.
		Stakeholder assessment	The company is conducting the official stakeholder assessment.
	Implementation process	Assessment	Assessments such as a product or company LCA or Ecovadis assessment, that companies have performed to gather data on their process.
		General CSRD	General steps undertaken in the CSRD implementation process.
		Small steps	Advice to split the implementation into small steps.
	Motivation	Digitalization	Motivation in terms of digitalization.
		Intrinsic	The motivation for CSRD compliance is from intrinsic nature.
		Obligation	The organization feels like CSRD is mainly an obligation.
		Opportunity	The organization is motivated because they experience that CSRD brings opportunities.

	Stakeholders	Attitude	The attitude of stakeholders.
		CSRD team	How the CSRD team within the organization is constructed.
		Engagement	How organizations engage with their stakeholders.
		Parties involved	The external parties that are involved in the process.
	Transparency		The importance to be transparent as a company, both on the way of reporting and the data.
	Vision	CSRD goals	The goals related to CSRD implementation in general.
		Digitalization	The vision for digitalization in a CSRD context.
		General sustainability	The general vision for sustainability.
Challenges	Data	Data diversity	Diversity in how the data is shared by stakeholders; both in format and the content of data.
		Data formats	Challenges related to the format of exports.
		Data hierarchy	Challenges related to the different data hierarchies within systems.
		Data quality	Data quality in terms of unreliable data.
		Data translation	The translation from data to information.
		Gaps in data	In these cases, data is not complete.
		Labour-intensive	Retrieving and checking data is labour-intensive.
		Manual work	Manual data gathering.
		System functions	Systems are not designed for the type of reporting that is now required and have different core functions.
	Internal	Capacity	Capacity in terms of human resources.
		Central direction	A lack of central direction poses a challenge for the CSRD implementation.
		Incentive	The challenge regarding incentive is how to get employees on board.
	External	Changing regulation	Changes regarding the CSRD regulation, making it difficult to keep up.
		Value chain	Challenges related to the value chain.
Metrics	Data availability		Comments regarding how much of the needed data for CSRD reporting is available, either internal or external.
	Reporting points		KPIs or metrics that the company is or will be reporting on.

9.8 ESSENCE OF THEMES

Data Infrastructure and Systems – red

This theme encompasses the technical backbone and tools necessary for data management. It includes aspects such as centralization of the landscape, the use of specific consolidation tools like data lakes and warehouses, the systems in place (existing and future software solutions), flexibility within and between these systems, and the broader motivation for digitalization. This theme also covers how the landscape is visualized, indicating the representation of data and applications.

CSRD Implementation – orange

This theme deals with the practical aspects of implementing CSRD requirements. It includes considerations like timelines, the complexity of the process, and the need for transparency. It also addresses the changes driven by insights gained during implementation and the importance of starting the process on time.

Data Points – grey

This theme focuses on the data that is retrieved and specific challenges that arise in doing so. It covers a range of issues from data availability and quality to manual work and data translation, capturing the operational aspects of handling data.

Stakeholder Involvement and Engagement – green

This theme focuses on the human and relational aspects, including the involvement and attitudes of stakeholders, the role of the CSRD team, and the engagement of various parties. It also covers internal challenges related to capacity, direction, and incentives.

Strategic Considerations – light blue

This theme aligns with the broader strategic considerations of the organization, linking data and sustainability initiatives such as CSRD to business value and overall vision. It additionally includes the central role that organizations can take as a strategic position.

Compliance Factors – blue

This theme covers regulatory and compliance-related aspects, such as the double materiality assessment as well as external challenges stemming from changing regulations and the broader value chain.

Drivers for Motivation – purple

This theme explores the different motivations behind adopting and implementing the CSRD directive. It includes intrinsic motivation, obligation, and opportunities.

9.9 DESIGN PROCESS

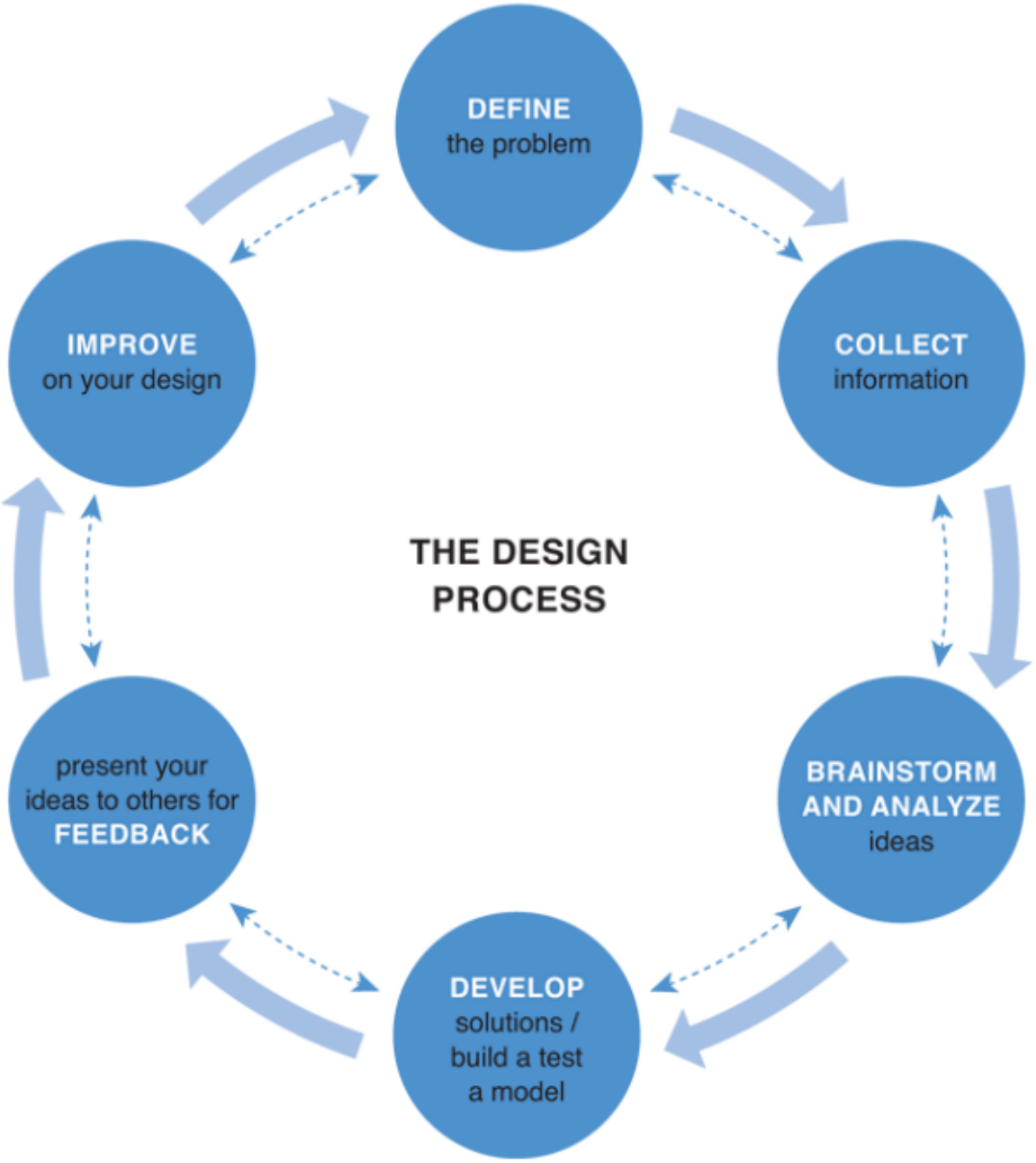


Figure 24 - Design Process by Munro (2016)

9.10 ROADMAP GUIDELINES

CSRD ROADMAP – the guidelines

Building support

- Engage and involve employees; communicate the value of CSRD
- Establish a cross-departmental reporting team
- Implement effective change management practices
- Standardize data exchange with stakeholders

Mapping the IT landscape

- Visualize the current IT landscape (as-is); use TOGAF & ArchiMate
- Incorporate external stakeholders in mapping
- Assess potential issues
- Visualize potential future situations (to-be)

Overview of metrics

- Identify required metrics based on the DMA
- Establish a "Definition of Done"; conduct a gap analysis
- Plan for future data quality improvements
- Visualize data process flows

Solution research

- Define criteria for solutions
- Explore available tools
- Address manual data collection challenges
- Create or update the to-be model

Centralizing data

- Select the appropriate technique (warehouse, lake, lakehouse)
- Address data dispersion
- Strive for a Single Source of Truth
- Standardize data and enhance data quality control

Realizing the to-be landscape

- Integrate selected tools with the centralized platform
- Initiate integrated reporting
- Establish ongoing evaluation processes and proactive monitoring
- Stay updated on technological advancements

9.11 EXPERT OPINION QUESTIONS

1. Effects: (artifact x context) produce effects?

- How does the roadmap help companies towards integrated CSRD reporting?
- To what extent can the roadmap be applied to projects focused on CSRD compliance?
- To what extent can the roadmap be tailored to specific projects focused on CSRD compliance across various industries?
- How can the roadmap be embedded in client projects?
- What organizational changes at Flawless are necessary to effectively embed the roadmap within future projects?

2. Requirements: do effects satisfy requirements?

- Does the roadmap provide clear and actionable guidelines that facilitate the transition towards integrated CSRD reporting?
- How does the roadmap add value to the work of consultants helping clients to achieve CSRD compliance?
- Does the roadmap add value for companies that embark on the path to CSRD compliance themselves?

3. Sensitivity: (artifact x alternative context) produce effects?

- In what ways can the roadmap be adapted for use in contexts other than projects focused solely on CSRD compliance (e.g., corporate strategy, sustainability reporting)?
- What assumptions need to be in place for the use of the roadmap in different contexts (e.g., varying organizational sizes and structures)?
- What adjustments, if any, are needed for the roadmap to remain relevant in rapidly changing regulatory environments?

4. Trade-offs: (alternative artifact x context) produce effects?

- What are other methodologies/tools (if any) that can facilitate integrated CSRD reporting?
- How does the roadmap perform as opposed to alternative artifacts (e.g., ease of implementation and adaptability)?

9.12 SURVEY QUESTIONS

A Roadmap Towards Integrated CSRD Reporting

You are being invited to participate in a research study titled a Roadmap Towards Integrated CSRD Reporting. This study is being done by Iris van Heijnsbergen from the Faculty of Electrical Engineering, Mathematics and Computer Science at the University of Twente.

The purpose of this research study is to gather feedback on the roadmap for CSRD reporting and will take you approximately 15 minutes to complete. The data will be used for the master thesis of the researcher.

Your participation in this study is entirely voluntary and you can withdraw at any time.

We believe there are no known risks associated with this research study; however, as with any online related activity the risk of a breach is always possible. To the best of our ability your answers in this study will remain confidential. We will minimize any risks by making the survey anonymous and storing the data only for a limited the time, till the graduation of the researcher in October 2024. Once data has been used in the research it will safely be deleted.

Study contact details for further information: Iris van Heijnsbergen, *i.m.vanheijnsbergen@student.utwente.nl*

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Secretary of the Ethics Committee/domain Humanities & Social Sciences of the Faculty of Behavioural, Management and Social Sciences at the University of Twente by ethicscommittee-hss@utwente.nl

Consent

1. Based on the information of the study, I consent to participate in this survey. *

   Answering "No" will make the survey jump to the end, skipping all questions.

   ◯ Yes

   ◯ No

Roadmap Effectiveness

2. How helpful do you find the roadmap and guidelines to guide your company in the process towards integrated CSRD reporting? *

   Here you can for instance consider how the steps and guidelines can support you in the process and whether they are applicable to a project in practice.

	Not helpful	Not very helpful	Neutral	Somewhat helpful	Very helpful
Scale	◯	◯	◯	◯	◯

3. How applicable do you think the roadmap is to projects focused on CSRD compliance? *

   E.g., the steps and guidelines incorporate important steps necessary for this process.

	Not applicable	Not very applicable	Neutral	Somewhat applicable	Highly applicable
Scale	◯	◯	◯	◯	◯

4. How easy would it be to integrate the roadmap into ongoing CSRD compliance projects? *

   Points to consider are whether the roadmap and guidelines align with current practices and whether it is possible to start using the roadmap from a different phase than the first one.

	Not easy	Not very easy	Neutral	Somewhat easy	Very easy
Scale	◯	◯	◯	◯	◯

5. How well do you think the roadmap can fit projects in different industries? *

   E.g., is the roadmap applicable to your own industry as well as other industries

	Not well at all	Not so well	Neutral	Somewhat	Very well
Scale	◯	◯	◯	◯	◯

6. Do you think your organization can effectively use the roadmap, considering current ongoing practices? *

   Explanation can be given at the next question.

   ◯ Yes

   ◯ No

7. What changes, if any, do you think your organization needs to make to effectively use the roadmap?

Enter your answer

Requirements

8. Do the roadmap and guidelines provide clear steps for transitioning to integrated CSRD reporting? *

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Scale	◯	◯	◯	◯	◯

9. How much value do the roadmap and guidelines add to your or consultants' work of achieving integrated CSRD reporting? *

E.g., the roadmap provides a helpful structure or valuable points that were not considered before

	No value	Little value	Neutral	Some value	A lot of value
Scale	◯	◯	◯	◯	◯

10. What, if any, are specific features that you consider as positive aspects of the roadmap and guidelines in their current state? *

Enter your answer

11. What, if any, are specific features that you miss in the roadmap or the guidelines? *

Enter your answer

Flexibility and adaptability

12. Do you consider the roadmap to be flexible enough to stay relevant in a rapidly changing regulatory environment like the CSRD? *

- ◯ Yes
- ◯ No

13. What conditions, if any, do you think are necessary for the roadmap to work in different types of organizations (e.g., size, structure)? *

Enter your answer

General remarks

14. Are there any additional remarks or points of feedback that you would like to share?

Enter your answer