Organisational Maturity Assessment during the Paradigm Shift from Monoliths to Data Mesh

Design Science Research in Developing a Data Mesh Maturity Assessment Model

Master thesis submitted to Delft University of Technology

in partial fulfilment of the requirements for the degree of

Master of Science in **Complex Systems Engineering and Management**

Faculty of Technology, Policy, and Management

by

C. Jonkman

4597915

To be defended in public on July 6th, 2023

Graduation Committee:

Chairperson:	Dr. ir. G.A. de Reuver	Section ICT
First Supervisor:	Dr. M.L.C. de Bruijne	Section O&G
Second Supervisor:	Dr. ir. G.A. de Reuver	Section ICT
External Supervisor:	M. van Andel	Accenture

TU Delft   accenture

Preface

I am delighted to share my thesis with you. With the completion of this research, I am concluding my Master's degree in Complex Systems Engineering and Management at Delft University of Technology. I would like to take this preface as an opportunity to share experiences from this research that represent the personal developments that have shaped me personally during my studies. In addition, I would like to express my appreciation to those who have contributed to and supported me over the past half a year.

From the very beginning, the research topic of data mesh has excited me due to its novelty and emergence in the field of data management. My choice of this topic exemplifies the enthusiasm, interest, and passion I have developed for data management during my studies. In addition, my research involved managing various interests, opinions, and dependencies. I actively sought out this dynamic in my research by conducting my master's thesis both at the university and during an internship at Accenture. Moreover, the choice of qualitative design science research, which entailed engaging with numerous individuals and external organisations, provided me with the opportunity to navigate these different perspectives. Although this position could pose challenges, I appreciate others' involvement, reflecting my preference for collaboration in my work. Furthermore, I would like to share that I consider my research to be a unique opportunity for personal development, in which my internship has played a significant role as I identified a strong practical relevance for my research. I approached the work results-oriented, aiming to deliver a valuable and ready-to-use deliverable for Accenture. During this process, I was privileged to access their international network of experts and external organisations. The experience of collaborating internationally, involving external leading organisations in my research, and receiving appreciation for my work has brought me a great sense of fulfilment throughout my research. Maira van Andel has supported me throughout this process, and I would like to thank her as my Accenture supervisor. I greatly appreciate her involvement in my research, critical perspective on my work, and assistance in connecting me with colleagues and organisations.

In addition to conducting my research for Accenture, Mark de Bruijne and Mark de Reuver have contributed to my academic development as university supervisors by continuously encouraging critical reflection on scientific relevance and methodology. I would like to express my gratitude to them for their support. Mark de Bruijne, I would like to thank you for your high level of involvement during the research and your detailed feedback. Your expertise in the field of organisation and governance has enabled me to critically reflect on the social dimensions of data mesh. Furthermore, I appreciated that you gave me freedom in my approach to conducting research and supported me throughout that process. Working towards deadlines was personally preferable for me as it helped to maintain focus on intermediate deliverables. However, I also acknowledge that it sometimes required us to reflect on matters that were still in progress. I would also like to thank Mark de Reuver for his critical and scientific perspective on my research. Moreover, the data sharing seminar you organised was valuable to attend as it allowed me to explore data mesh from the perspective of cross-organisational data sharing. At last, I would like to express my gratitude to everyone else who has made valuable contributions to my research and provided support throughout the research process.

Dear reader, I hope you will enjoy reading my thesis and that this study will continue to be of great value to you in the future.

Christian Jonkman

July 6th, 2023

Executive Summary

Incorporating big data into decision-making provides a substantial competitive advantage, leading organisations to increasingly adopt a data-driven strategy. However, the adoption by organisations often remains unsuccessful due to limitations associated with monolithic data architectures, such as data lakes and data warehouses. Data mesh is introduced as a decentralised socio-technical approach to alternatively manage data, aiming to overcome the limitations and gain the benefits of embracing a data-driven strategy. However, there is a lack of guidance on how to implement data mesh. The availability of generic and concrete data mesh implementation steps, including a maturity assessment, would be helpful for organisations. Consequently, this research proposes the design of a Data Mesh Maturity Assessment Model (DMMAM). In response to the main research question: *"How to assess the maturity of a data mesh implementation within an organisation?"*, enabling the assessment of how mature a data mesh implementation is, by means of the DMMAM, would provide the guidance that is currently lacking for organisations. The qualitative Design Science Research Methodology is employed to structure the design process. Literature research, interviews, and cases are conducted to explore the contribution of, as well as design, demonstrate, and evaluate the DMMAM.

This research shows that the developed DMMAM evaluates data mesh based on four maturity levels, classified as Level 0: *Non-Initiated*, Level 1: *Conceptual*, Level 2: *Defined*, and Level 3: *Achieved*, and that data mesh is represented by five dimensions: A. *Data Foundation & Organisational Change*, B. *Domain Oriented Decentralised Data Ownership & Architecture*, C. *Data as a Product*, D. *Self-Serve Data Infrastructure as a Platform*, and E. *Federated Computational Governance*. These five dimensions are collectively represented by 54 characteristics. For each characteristic, labels for the People, Process, Technology (PPT) perspectives are assigned. Additionally, questions are formulated, and criteria and requirements are provided for all characteristics at each maturity level. This enables participants to self-assess their organisation's maturity by individually rating 54 questions based on the current and target levels. Conducting the self-assessment yields various outcomes, including an overall data mesh maturity score, individual dimensional maturity scores, and maturity scores from PPT-perspectives. Moreover, the assessment helps to identify maturity gaps and allows benchmarking to compare results across organisations, providing organisations with guidance for improvement. The demonstration and evaluation of the DMMAM through maturity assessments for three organisations have demonstrated its applicability and usefulness. However, it is important to acknowledge that this research represents the first attempt to provide a comprehensive framework for assessing data mesh maturity in organisations and is not without limitations.

Future research is proposed to further refine and improve the DMMAM, supported by data mesh SME's and practitioners, to ensure that the model remains up-to-date with the latest available research on data mesh. In addition, including additional guidance as an outcome of the maturity assessment would make the assessment more actionable and pragmatic. Furthermore, examining the optimal assessment structure will enhance the model's reliability and validity. Moreover, expanding the benchmark functionality will enable statistical generalisations and comparisons for organisations within and across industries. At last, it is suggested to do further research about examining the overall contribution of data mesh as a strategy element towards becoming data-driven.

Keywords: *Big Data* · *Data-Driven* · *Data Mesh* · *Maturity Assessment Model* · *Design Science Research*

Nomenclature

List of Abbreviations

$3\,V's$	Volume Velocity Variety
API	Application Programming Interface
$BDMAM$	Big Data Maturity Assessment Model
BI	Business Intelligence
CDO	Chief Data Officer
CEO	Chief Executive Officer
CIO	Chief Information Officer
CMM	Capability Maturity Model
$CoSEM$	Complex Systems Engineering and Management
CTO	Chief Technology Officer
$DAMAM$	Data Analytics Maturity Assessment Model
$DMMAM$	Data Mesh Maturity Assessment Model
$DMMM$	Data Management Maturity Model
DQ	Data Quality
$DSRM$	Design Science Research Methodology
ETL	Extract Transform Load
$GDPR$	General Data Protection Regulation
ID	Identity
IT	Information Technology
KPI	Key Performance Indicator
$MECE$	Mutually Exclusive Collectively Exhaustive
MSc	Master of Science
NA	Not Available
PLM	Product Life-cycle Management
PPT	People Process Technology
ROI	Return On Investment
SLO	Service Level Objective
SLR	Systematic Literature Review
SME	Subject Matter Expert

Contents

List of Figures

List of Tables

Phase A

Problem Identification and Motivation

1

Introduction

1.1. Problem Introduction

1.1.1. Context

The quantity of available data and data use has grown exponentially in recent years (Naga Rama Devi, 2019; Singh et al., 2022). This data generated at an exponential rate has given rise to the concept of big data (Rawat & Yadav, 2021). The potential benefits associated with big data have attracted the large attention of organisations (Surbakti, 2020). According to Surbakti, leading organisations in the future are those who could effectively use big data. Big data has therefore become a significant asset for organisations (Al-Sai et al., 2019).

Effectively incorporating big data into decision-making offers a substantial competitive advantage to both public and private organisations (Cuenca et al., 2021; Hupperz et al., 2021). Being intelligently empowered as an organisation has the benefit of providing a better customer experience, evaluating new business opportunities, and reducing operational costs (Faizi et al., 2017; Rivera & González, 2022; Tariq et al., 2021). In addition to the potential benefits for organisations, also customers take advantage (Jaiswal & Bagale, 2017). To illustrate, operational excellence through data utilisation could reduce customer costs and providing a better customer experience improves customer retention, satisfaction, and overall service quality (Chehri et al., 2022; Holzer & Karkoschka, 2019; Strengholt, 2023; J. Yin et al., 2016). Due to these benefits, Berndtsson et al. (2018) state that more often organisations adopt a strategy of becoming data-driven since data-driven organisations are more likely to become successful compared to organisations relying on intuition (Berndtsson et al., 2020).

However, the number of organisations which successfully transform into data-driven organisations remains low due to organisational and technical challenges (Svensson & Taghavianfar, 2020). These challenges are often linked to the limitations demonstrated by the traditional monolithic data architectures, such as data lakes and data warehouses (Machado et al., 2021, 2022). These data architecture types do not satisfy the growing needs of organisations. According to Machado et al. (2022), bottlenecks are often associated with the centrally organised data teams. Additionally, there is a lack of alignment between organisational needs and the technical architecture implemented.

To illustrate, in an organisation where data is produced and needed by everyone, but without clear data ownership, problems regarding data quality are likely to arise. As a result, this could adversely impact the value of analytical data. According to Dehghani (2022a), a paradigm shift is needed to overcome these limitations.

Dehghani presents data mesh as a decentralised socio-technical approach to managing data which considers domains as the primary concern, employs platform thinking, conceptualises data as a product, and introduces federated decision-making. More specifically, it is suggested that this approach is the convergence of four main dimensions, which refer to *domain oriented decentralised data ownership and architecture*, *data as a product*, *self-serve data infrastructure as a platform*, and *federated computational governance*. In a data mesh, these dimensions shape the decentralisation of the organisational structure and technical architecture. This research will examine and discuss this novel approach, including the provision of a detailed definition in Chapter 2: Theoretical Background.

Dehghani (2019) coined data mesh when she worked for over twenty years at the global technology company Thoughtworks in North America as a principal consultant and as a member of the technology advisory board. The article in which she introduced data mesh, titled *How to Move Beyond a Monolithic Data Lake to a Distributed Data Mesh*, was supported by Martin Fowler, a British software developer and author of various books and articles with a strong reputation among practitioners, academics, and researchers (Fowler, 2000). Since the introduction of data mesh, the progressive line on *Google Trends* for the search term *Data Mesh*, as highlighted by Goedegebuure et al. (2023), reflects the surging popularity of the approach. Moreover, Miner et al. (2023) state that data mesh is currently perceived as one of the top trends in analytics and business intelligence (BI). Furthermore, Dehghani published in the past two years the books *Software Architecture: The Hard Parts* and *Data Mesh* on O'Reilly, which is for over 40 years a widely recognised platform to learn about future technology (O'Reilly, 1978). As a result of the increased recognition, Dehghani is acknowledged globally by practitioners and researchers as the founder of data mesh.

The introduction of data mesh could have been expected state Ford et al. (2021) and Hechler et al. (2023), as it is an outstanding example of the ongoing incremental evolution observed in organisational information management. Ford et al. argue that the introduction of new capabilities brings forth new perspectives, which in turn help address enduring challenges from the past, such as the separation of domains from data, resulting in unclear data ownership. However, due to the novelty of data mesh, it is a long way ahead to argue that data mesh will definitively overcome the limitations of traditional data architectures and whether it contributes to becoming data-driven (Driessen et al., 2023). Moreover, criticism has already been expressed about data mesh, as stated in the book *Data Management at Scale* from Strengholt, published by O'Reilly. Strengholt states that the description of data mesh by Dehghani is considered incomplete because it lacks guidance in areas such as data reusability, data accessibility, and master data management. Nevertheless, numerous organisations are currently contemplating the adoption of data mesh to overcome the limitations of traditional data architectures (Panigrahy et al., 2023). Machado et al. (2022) state that adopting data mesh as a data strategy element would foster the process for organisations to become data-driven.

1.1.2. CoSEM Research

This research is performed as part of the programme *Master of Science* (MSc) in *Complex Systems Engineering and Management* (CoSEM). Exemplary CoSEM-studies focus on solving complex socio-technical problems with a technical, institutional, and process component while taking into consideration both public and business values. The programme's aim is to design and assess the impact of socio-technical solutions in organisations which contain systems engineering approaches, effective management strategies, and ethical aspects.

This research focuses on data mesh. Dehghani (2022a) presents data mesh as a decentralised socio-technical approach to sourcing, managing, and accessing data. Moreover, data mesh is introduced to focus on organisations operating in complex and large-scale environments. As mentioned, data mesh encompasses four main dimensions, namely *domain oriented decentralised data ownership and architecture*, *data as a product*, *self-serve data infrastructure as a platform*, and *federated computational governance*.

Aligned with the CoSEM-programme, data mesh encompasses technical, institutional, and process components, which will be explained through the four dimensions and the presentation of six dimensions of change, introduced by Dehghani. The technical artefacts of data mesh relate to the design of a decentralised data architecture, the design of data products as architectural units, the establishment of a self-serve data infrastructure, and embedding data governance policies in the data products and data infrastructure to enable automation. The institutional artefacts pertain to the data governance operating model which in a data mesh balances global agreement with federated decision-making. The data governance operating model impacts how an organisation operates around its data and helps establish an organisational governance structure (Brennan et al., 2018). For instance, the data governance operating model outlines how the organisation defines roles and responsibilities, organisational terms, and domain types. Global agreement refers to overarching policies that address guiding values, incentives, and the composition of federated domain teams. Computational governance entails that policies need to be automated, which simultaneously involves ethical considerations. The process artefacts are included by perceiving the adoption of data mesh as a paradigm shift, which involves both technical and organisational changes. Dehghani presents six shifts, which are the transition from a centralised towards decentralised data ownership, from collecting data in monolithic data architectures towards connecting data through a distributed mesh of data products, from data as a byproduct of code towards data and code as one unit, from top-down towards federated computational governance, from data as an asset to be collected towards data as a product to share and connect, and from a fragmented towards a well-integrated infrastructure. These six shifts show how data mesh involves various technical and organisational changes in relation to traditional data architectures. Furthermore, numerous stakeholders within an organisation are engaged during these shifts, including data owners, data consumers, analytical data teams, governance teams, and data infrastructure teams (Krystek et al., 2023). The shifts also impact team structures, functional roles, and new responsibilities (Voß, 2022). Altogether, this research covers the various perspectives which are exemplary for the CoSEM-programme.

1.1.3. Societal Relevance

The societal relevance of this research relates to the contribution it aims to make to help organisations in transforming towards becoming data-driven. Being data-driven presents various benefits for organisations and also has a positive impact on their customers. This research looks therefore into data mesh as a decentralised socio-technical approach which could help organisations in realising the value of being data-driven.

1.2. Literature Review

1.2.1. Scientific Relevance

The scientific relevance of this research refers to the aim to contribute to existing scientific work by researching a knowledge gap in the literature. To search for existing scientific work related to data mesh, the Scopus database was used. The search term *Data Mesh* was applied within *Article Title, Abstract, and Keywords*. This search resulted in 79 documents. Since the concept of data mesh was coined in 2019, the search was limited to only documents from 2019 and beyond, which resulted in 29 documents. For the remaining 29 documents, the abstract was analysed. In the end, 10 articles seemed relevant to this literature review. The other 19 documents only implicitly address data mesh. The relatively low number of available scientific publications reflects the novelty of the approach.

In order to identify a knowledge gap, the selected documents will be analysed and structured. The literature will be analysed in terms of stating what already has been researched in existing publications. The structure will be provided by looking into different phases of the implementation process. This approach is chosen because the adoption and application of data mesh, considered as a transition from a traditional data architecture, could be seen as an implementation of its dimensions, states Dehghani. These dimensions represent the organisational and technical components that data mesh entails. To elaborate on this in more detail, publications by Aarons et al. (2011), Blase & Fixsen (2013), Fixsen et al. (2005), and Fixsen et al. (2009), which examine the literature on implementations, will be used to define *Implementation* for the purpose of this research. Fixsen et al. (2005, p. 11) define implementation as "*... a specified set of activities designed to put into practice an activity or program of known dimensions.*" It will be suggested that, in the context of data mesh, implementation refers to the activities involved in deploying the data mesh dimensions, represented by its components. Blase & Fixsen (2013, p. 2) define *Components* as "*... the essential functions or principles, and associated elements and intervention activities that are judged necessary to produce desired outcomes.*" Furthermore, Fixsen et al. (2009) suggest that research on the implementation of these components extends across different phases. They argue the importance of this distinction to more rapidly advance research and practice in the field of implementations. Therefore, existing publications related to data mesh will be structured based on implementation phases. Since no prescriptive implementation approach exists for data mesh due to its novelty, the EPIS-framework presented by Aarons et al. will be used since it highlights key phases of a general implementation process. Four consecutive phases are described by the framework, which are *exploration*, *preparation*, *implementation*, and *sustainment*.

The *exploration phase* evaluates the needs and potential for implementation. First of all, this has been researched by Machado et al. (2021) and Scrocca & Tommasini (2021) by looking into the usefulness and benefits of data mesh, providing a motivation for its appearance and features, and exploring the emergence of a domain-driven data design. Secondly, problems and limitations from traditional data architectures were approached from a data mesh perspective to obtain an understanding of the differences (Mehmandarov et al., 2021; Priebe et al., 2021). Thirdly, a holistic view of data governance in data mesh infrastructures is discussed to explore how organisations could agree upon overarching rules and regulations (Joshi et al., 2021). Fourthly, the potential of data mesh is discussed in smart cities (Soe et al., 2022). The *preparation phase* refers to how the approach could be adopted by an organisation. Hooshmand et al. (2022) and Machado et al. (2021) discuss this adoption by proposing an approach for the transformation process towards the data mesh implementation. The *implementation phase* addresses how to put the approach into place. Research has been performed by Machado et al. (2022) concerning a proposal on how the technical architecture of a data mesh could be implemented. Once the approach is embedded, it is important to evaluate it in the *sustainment phase* for ongoing monitoring and quality assurance. In addition, Aarons et al. state that lessons learnt from previous phases may have an impact on future implementation efforts for sustainment. Evaluation has been conducted in the article from Joshi et al. by a case study related to addressing data governance challenges. In addition, issues that emerge in guaranteeing the privacy of distributed mesh data are examined (Podlesny et al., 2022). Lastly, findings are reported from a Norwegian public sector organisation which is currently adopting a data mesh (Vestues et al., 2022). Analysing these articles shows that most literature is focused on the *exploration* of the approach. In the context of data mesh, exploration refers to gaining an initial understanding of the approach. This includes exploring the benefits and challenges, examining its features, and evaluating its potential in addressing the limitations posed by traditional data architectures. To contribute to what has not yet been researched, obtaining a better understanding of how a data mesh could be *prepared* and *implemented* would add scientific value.

Future research is proposed by Machado et al. (2021) and Vestues et al. about a detailed approach, consisting of rigorous and concrete steps, for the design and implementation of data mesh. In addition, it is mentioned by Machado et al. (2022) that it would be valuable if practitioners have a starting point for the implementation. So, future research is proposed about explicit steps to be taken, from a starting point towards an endpoint, to implement data mesh.

Literature also showed that there is currently no generic approach, implying not being specific to any particular organisation or industry, for implementation. Machado et al. (2022) note that there are several technologies suitable for implementing the data mesh components. Vestues et al. mention the varying views by practitioners and researchers on how an organisation could implement data mesh. In the article from Joshi et al., case-specific factors are identified to find a fit-for-purpose approach to implement the data mesh governance. Additionally, it was mentioned that the evaluation of organisational maturity is crucial to gain clarity on which data mesh aspects need to be implemented and which already have been implemented. These findings build on the statement that no path from an immature towards a mature data mesh implementation is yet available. Consequently, there is a need for a generic approach with a maturity assessment, which could guide organisations during the data mesh implementation process.

Combining the need for generic and concrete data mesh implementation steps, including a maturity assessment, the design of a Data Mesh Maturity Assessment Model (DMMAM) is proposed to fill the knowledge gap. The DMMAM could be used to assess whether or not data mesh components have already been implemented in an organisation, and if so, to determine their level of maturity. From this starting point, the DMMAM shows how higher maturity levels could be achieved by implementing or further enhancing the data mesh components. The DMMAM presents, in terms of characteristics and maturity levels, a generalised approach to concrete steps to be taken to adopt data mesh.

1.2.2. Scientific Problem Statement

Berndtsson et al. state that organisations aim to transform towards becoming data-driven to obtain its benefits. However, according to Svensson & Taghavianfar, the number of organisations which successfully transform into data-driven organisations remains low due to the limitations posed by the monolithic data architectures. Data mesh is introduced by Dehghani as a decentralised socio-technical approach which tries to overcome these limitations. Despite the potential of data mesh, Section 1.2.1 addressed the current lack of guidance for an organisation on how to transform the data architecture approach towards a data mesh. This research argues that guidance would help organisations adopt data mesh and thereby contribute to their aim towards realising the value of being data-driven.

In the current situation, wherein guidance is lacking, organisations find themselves without generic and concrete steps to be taken to adopt data mesh. Consequently, organisations will still encounter the limitations as mentioned by Machado et al. (2021), which are inherent to the monolithic data architectures. As a result, these organisations are unable to effectively undergo the transformation, thus missing out on the benefits that could otherwise be attained.

To provide guidance to organisations on how to transform towards a data mesh, this research proposes the design of a DMMAM. This research will elaborate on how this model will provide the guidance that organisations need during the data mesh implementation process. Consequently, organisations will be guided as they take steps forward in overcoming the limitations of monolithic data architectures, thereby aiming to obtain the benefits of being data-driven.

Based on this scientific problem statement, the main research question, main research objective, and main research deliverable are formulated and provided in Figure 1.1.

Main Research Question

How to assess the maturity of a data mesh implementation within an organisation?

Main Research Objective

Enabling the assessment of how mature a data mesh implementation is within an organisation.

Main Research Deliverable

Design of a data mesh maturity assessment model.

Figure 1.1: Main Research Question, Objective, and Deliverable

This research seeks to answer the main research question referring to how the maturity of a data mesh implementation within an organisation could be assessed. Through the design of a DMMAM, the assessment of how mature a data mesh implementation is will be enabled. Organisations would obtain an understanding of whether or not, and to what extent, data mesh components have already been implemented. The main deliverable of this research is the design of a DMMAM, which presents, in terms of characteristics and maturity levels, a generalised approach on which concrete steps need to be taken to adopt data mesh. Assessing the organisational maturity provides guidance to support the data mesh implementation process, starting from the current maturity level and progressing towards a desired state.

The scope of this research focuses on data mesh within an organisation. The intra-organisational focus has been chosen because the scientific publications, as presented in Section 1.2.1, also perceived data mesh as an intra-organisational approach. Furthermore, perceiving data mesh as an inter-organisational approach would make the assessment dependent on other organisations. Correia et al. (2023) mention that most of the existing maturity assessment models do not account for the involvement of inter-organisational stakeholders. Moreover, Frick et al. (2013) state that maturity assessments for inter-organisational approaches are often inherent to contradictions. As a result, Frick et al. state that assessment models would become complex and therefore may limit the interest of users and could produce misleading results.

1.3. Research Design

1.3.1. Research Methodology

A qualitative design approach will be employed to provide the answer to the main research question. The research process will be structured by the *Design Science Research Methodology* (DSRM). According to Johannesson & Perjons (2014), the DSRM enables the creation of an artefact in the form of a model that supports developments in the area of information systems. This would imply the creation of a model focusing on the data mesh implementation. Moreover, Lasrado et al. (2015) state that the DSRM is widely adopted for developing maturity assessment models. Therefore, this methodology will be used to provide structure in this research for developing the DMMAM. The DSRM-framework provided by Peffers et al. (2007) will be used to carry out the research. This framework describes six design phases, which are the following:

A. Problem Identification and Motivation
B. Objective for a Solution
C. Design and Development
D. Demonstration
E. Evaluation
F. Communication

The design phases will structure the main process of the research design. However, the *communication* phase will not be included in this research. This phase would focus on the communication of the importance of the designed model after development to other researchers and professionals. In terms of this research, it does not directly contribute to answering the main research question and is therefore left out.

For the design phases, sub-questions are formulated which each have their corresponding sub-objective and sub-deliverable. According to Peffers et al., the *problem identification and motivation* phase defines the specific research problem and justifies the value of the solution. Since this phase rather focuses on the research definition than contributing towards answering the main research question, it is included as the introduction and as theoretical background in respectively Chapter 1 and 2. This also means this phase does not address a sub-question.

1.3.2. Research Sub-Questions

Phase B. *Objective for a Solution*

The four remaining design phases all correspond to a formulated sub-question. To begin with the *objective for a solution* phase, Peffers et al. state that this phase infers the desirability for guidance as objective from the *problem identification and motivation* phase, which will be the DMMAM. This model would help organisations in the transformation towards a data mesh. Therefore, exploring how the DMMAM helps organisations implement data mesh is the sub-objective in this phase. To obtain an understanding of the relevance of the DMMAM, literature research will be performed. The literature will be examined to uncover the contribution of using this model to the data mesh implementation process. Given the sub-objective and sub-deliverable of Phase B, Figure 1.2 also presents the corresponding sub-question.

Sub-Question 1

How does a data mesh maturity assessment model contribute to the data mesh implementation process?

Sub-Objective 1

Explore how the data mesh maturity assessment model helps organisations implement data mesh.

Sub-Deliverable 1

Relevance of a data mesh maturity assessment model to the implementation process.

Figure 1.2: Design Phase B: *Objective for a Solution*

Phase C. *Design and Development*

The *design and development* phase has the sub-objective of designing the DMMAM. The deliverable in this phase is a developed DMMAM which enables the assessment of how mature an organisation's data mesh implementation is. For answering the sub-question about what model could be designed, literature research and expert interviews will be conducted. Literature research will present the contribution of a systematic engineering design approach in addressing design considerations. Expert interviews will be carried out in an informal semi-structured manner to discuss the design choices. Figure 1.3 presents the sub-question, sub-objective, and sub-deliverable for Phase C.

Sub-Question 2

What model could be designed to assess the maturity of a data mesh implementation within an organisation?

Sub-Objective 2

Designing the data mesh maturity assessment model.

Sub-Deliverable 2

Developed data mesh maturity assessment model.

Figure 1.3: Design Phase C: *Design and Development*

The sub-question from Phase C consists of three sub-sub-questions. Figure 1.4 presents the sub-sub-question about how to design the DMMAM. To answer this question, a systematic engineering design approach will be introduced.

Sub-Sub-Question 2.1

How to design the data mesh maturity assessment model?

Sub-Sub-Objective 2.1

Defining a systematic approach to design the data mesh maturity assessment model.

Sub-Sub-Deliverable 2.1

Systematic approach for the design of the data mesh maturity assessment model.

Figure 1.4: Design Phase C: *Design and Development: Systematic Approach*

Figure 1.5 provides the sub-sub-question about what the design of the DMMAM will be. By following the systematic engineering design approach, design considerations will be motivated by insights gathered from the expert interviews.

Sub-Sub-Question 2.2

What will be the design of the data mesh maturity assessment model?

Sub-Sub-Objective 2.2

Selecting a preferred data mesh maturity assessment model design.

Sub-Sub-Deliverable 2.2

Preferred data mesh maturity assessment model design.

Figure 1.5: Design Phase C: *Design and Development: Preferred Design*

Figure 1.6 shows the sub-sub-question about what outcomes will be provided by the DMMAM.

Sub-Sub-Question 2.3

What outcomes could be provided by using the data mesh maturity assessment model?

Sub-Sub-Objective 2.3

Defining the outcomes of the data mesh maturity assessment model.

Sub-Sub-Deliverable 2.3

Outcomes of the data mesh maturity assessment model.

Figure 1.6: Design Phase C: *Design and Development: Model Outcomes*

Phase D. *Demonstration*

In the *demonstration* phase, the proof-of-concept DMMAM will be demonstrated in cases. Cases will be organised in which multiple organisations will be involved to conduct a maturity assessment by using the developed DMMAM. By using this model in cases, the applicability and usefulness of the DMMAM will be observed, which is the sub-objective in this phase. The decision to consider applicability and usefulness is based on the work from Gökalp et al. (2022), in which these concepts were used for the demonstration and evaluation of their data maturity assessment model. Moreover, this work is also expected to be useful in this research as they subsequently present criteria which will be used in Phase E to evaluate the extent of the DMMAM's applicability and usefulness, in terms of reliability. While applicability refers to the practical feasibility of demonstrating the DMMAM, usefulness pertains to the ability to provide the desired outcomes as will be defined in Phase C. The sub-deliverable in this phase will be a demonstrated DMMAM. To summarise, Figure 1.7 shows all sub-elements from Phase D.

Sub-Question 3

How to demonstrate the data mesh maturity assessment model?

Sub-Objective 3

Observing the applicability and usefulness of the data mesh maturity assessment model.

Sub-Deliverable 3

Demonstrated data mesh maturity assessment model.

Figure 1.7: Design Phase D: *Demonstration*

Phase E. *Evaluation*

The last sub-question focuses on evaluating to what extent the designed DMMAM is applicable and useful during the cases. In contrast to Phase D, which examines whether the DMMAM is applicable and useful at all, Phase E assesses the extent to which it is applicable and useful. This also implies that reflections on how the cases have progressed will be evaluated in Phase E, rather than in Phase D. In this *evaluation* phase, it will be aimed to conclude the extent of applicability and usefulness, in terms of looking into the reliability and validity of the model. The sub-deliverable will be the final DMMAM. All sub-elements from Phase E are presented in Figure 1.8.

Sub-Question 4

To what extent is the designed data mesh maturity assessment model applicable and useful?

Sub-Objective 4

Concluding about the extent of applicability and usefulness of the data mesh maturity assessment model.

Sub-Deliverable 4

Final data mesh maturity assessment model.

Figure 1.8: Design Phase E: *Evaluation*

1.4. Research Structure

The overall structure of this thesis report takes the form of five phases. Phase A focuses on formulating the research definition. The research definition will be introduced in Chapter 1 and the theoretical background will be provided in Chapter 2. Phase B continues in Chapter 3 by explaining the contribution of the DMMAM to the data mesh implementation process. In Phase C, the developed DMMAM will be presented. This phase will address the systematic engineering design approach in Chapter 4, the actual model design in Chapter 5, and will elaborate in Chapter 6 on which outcomes could be provided by the model. Phase D aims to deliver a demonstrated model by using the designed DMMAM in cases, which will be discussed in Chapter 7. At last, Phase E evaluates the extent of the DMMAM's applicability and usefulness in Chapter 8. The conclusion, discussion, and recommendations will be given in respectively Chapter 9, 10, and 11. At last, Chapter 12 ends with the research reflection.

A visualisation of the complete research design is given in Figure 1.9. This figure presents the research flow diagram consisting of the sequential DSRM-phases. For each phase where applicable, the sub-question is visualised together with the corresponding research method(s). Additionally, the figure shows which sub-objective is linked to each research phase. Finally, the sub-questions provide intermediate sub-deliverables. These are linked by outgoing arrows from the research phase towards the sub-deliverable.

The sub-question from Phase C is divided into three sub-sub-questions to progress into Phase D. Figure 1.10 presents a more detailed research flow diagram for only Phase C.

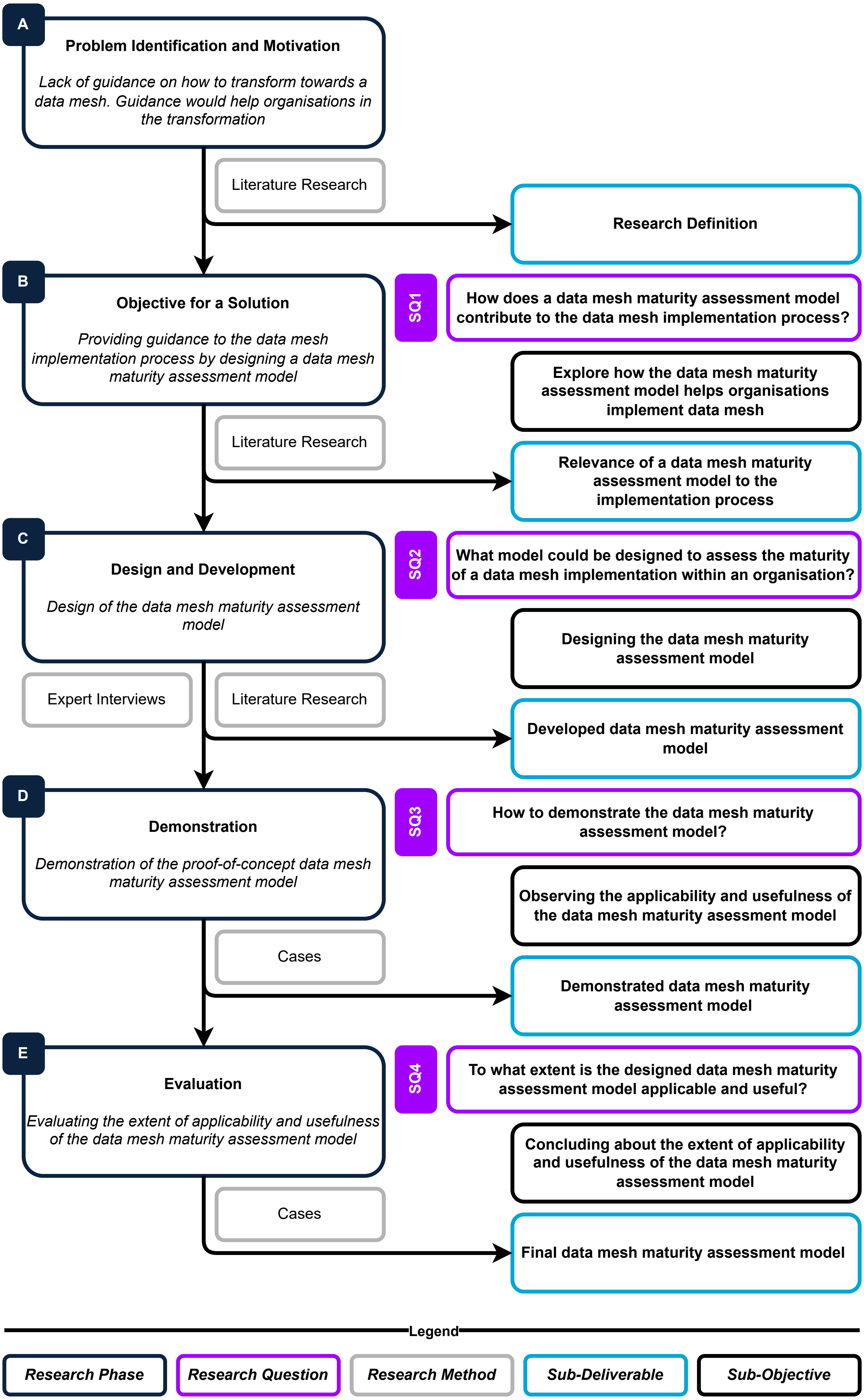


Figure 1.9: Research Flow Diagram

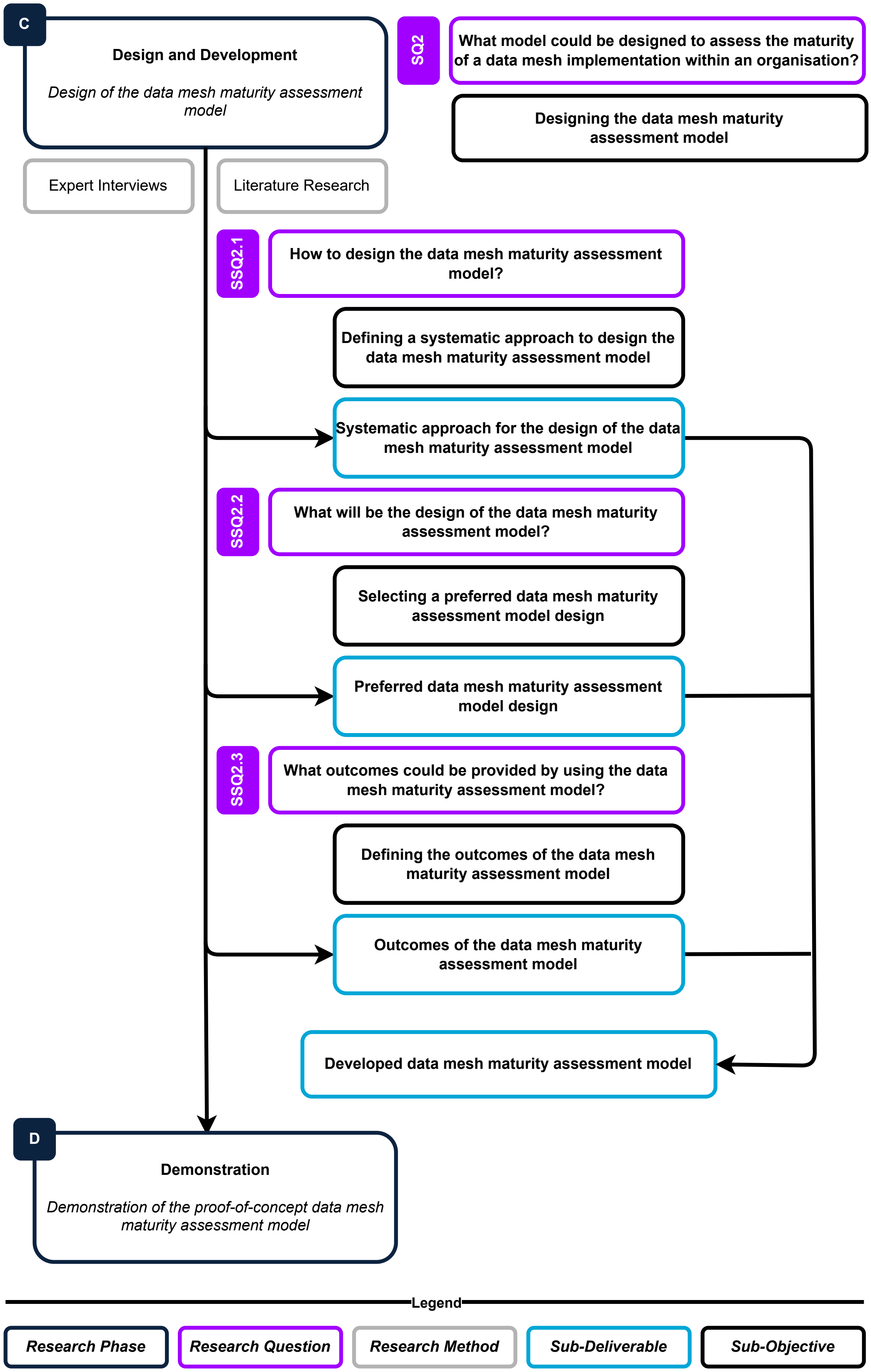


Figure 1.10: Research Flow Diagram: Design Phase C: *Design and Development*

2

Theoretical Background

Several core concepts in this research need to be defined beforehand for having a common understanding. This chapter will therefore provide an explanation of these concepts as theoretical background. This chapter is structured in two sections. Section 2.1 provides the definitions of the relevant concepts in this research. The outcomes of data mesh will be discussed in Section 2.2.

2.1. Defining Data Concepts

The term *Data*, and its different types as *Big Data*, *Operational Data*, and *Analytical Data* are often defined differently in literature. Therefore, Section 2.1.1 explains which definitions will be used in this research. Section 2.1.2 explains what *Data-Driven* entails for an organisation. *Data Architectures* and the relevant types in the context of this research will be discussed in Section 2.1.3. Lastly, Section 2.1.4 describes data mesh as a *Paradigm Shift*.

2.1.1. Defining Data

To define data, the definition provided by DAMA International (2017) will be used. DAMA International is a non-profit and vendor-independent association of professionals dedicated to advancing data management principles. According to DAMA International (1980), the association published the book *Data Management Body of Knowledge* to establish a shared terminology for data management, which is nowadays globally perceived as the dictionary for data management by IT professionals, executives, educators, and researchers. Due to this widespread recognition, several of their definitions are selected for this research. DAMA International (2017, p. 18) defines *Data* as "*... the information that has been stored in digital form.*"

Big Data - De Mauro et al. (2016, p. 131) identified and described primary research fields related to big data and proposed a more solid definition. *Big Data* is described as "*... the information asset characterised by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into value.*" Al-Mekhlal & Khwaja (2019) agree on De Mauro et al. that the main characteristics of big data are *variety*, *volume*, and *velocity*, also referred to as the 3 V's. Al-Mekhlal & Khwaja mention that variety refers to the different forms of data, such as structured, semi-structured, unstructured, and raw data. Volume refers to the large quantities of generated data. Velocity is included in the definition since big data moves at high speed from various sources.

Operational Data - The selected definitions for operational data and analytical data are provided by Dehghani (2022a), since she describes these different types of data in the context of data mesh. Dehghani (2022a, chap. 1, para. 12) defines *Operational Data* as the data which is used "*... to run the business and serve the end user*" and "*... sits in databases of microservices, applications, or systems of records that support business capabilities.*" In short, operational data refers to the information related to the internal functions of an organisation.

Analytical Data - Dehghani refers to analytical data as a fundamental ingredient for organisations to transition from intuition towards decision-making informed by observations and data-driven predictions. Whereas operational data refers to the data that supports running the business, *Analytical Data* is defined by Dehghani (2022a, chap. 1, para. 13) as "*... the historical, integrated, and aggregate view of data created as the byproduct of running the business.*" Analytical data serves data analysts and data scientists who perform statistical, diagnostic, or predictive analyses; create reports and visualisations; and train machine learning (ML) models. These analyses, insights, and models help to optimise organisational processes and improve organisations with intelligent automation.

Figure 2.1 visualises the relationship between *Analytics*, *Operations*, and *Big Data*. From the operations, operational data is collected, which is characterised by the 3 V's. Subsequently, from this large amount of information, known as big data, specific technology and analytical methods are used to transform this information into relevant insights. These insights could be used to improve the organisational processes (Lavasani et al., 2021; Marcinkowski & Gawin, 2021). According to McAfee & Brynjolfsson (2012), organisations which integrate big data and analytics into their operational processes are having higher productivity and profitability compared to their competitors. These empirically significant findings were obtained by comparing data-driven organisations with their industry competitors which did not embrace data-driven decision-making.

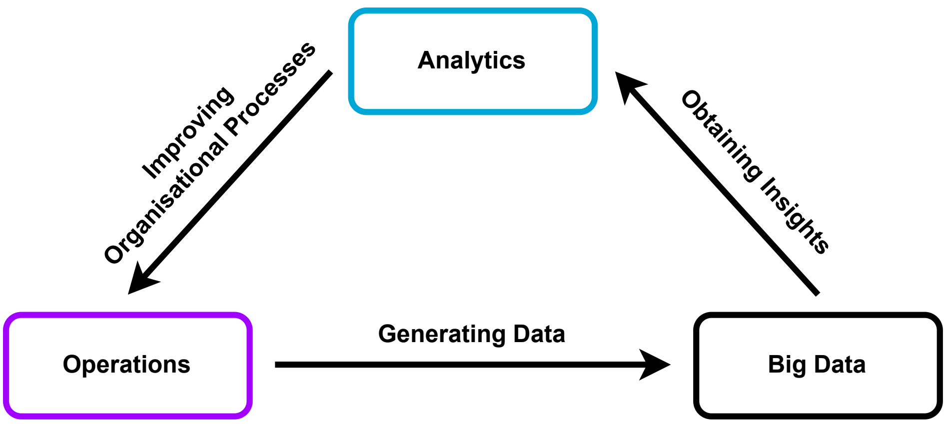


Figure 2.1: Operations, Big Data, and Analytics

2.1.2. Defining Data-Driven

Section 1.1.1 presented that data-driven organisations have a substantial competitive advantage compared to organisations relying on intuition (Berndtsson et al., 2020; Cuenca et al., 2021; Hupperz et al., 2021). Data mesh is introduced by Dehghani as data strategy element to become data-driven (Machado et al., 2022). To have a common understanding of what data-driven entails for an organisation, the term will be defined. The definitions provided by Anderson (2015) and Treder (2019) are selected as their books, titled respectively *Creating a Data-Driven Organization* and *Becoming a Data-Driven Organisation*, explore the term in depth. Moreover, these books have been retrieved from *O'Reilly* and *Springer*, which are both publishers with wide recognition in the field of technology. According to Treder, an organisation could be deemed as data-driven as soon as data is accepted across all levels of the organisation as a contributor to decision-making. Anderson explains contributing to decision-making as the active participation of a data-driven organisation in continuous testing to enhance operations and services, embracing a continuous improvement mindset, and utilising predictive modelling. Anderson adds to the definition from Treder that an organisation must have a data culture that acts on data with data processes in place to enable this decision-making. Consequently, this research considers an organisation to be *Data-Driven* as the organisation has accepted the use of data at every level as a contributor to enabling decision-making supported by a data culture and data processes.

2.1.3. Defining Data Architectures

As described by Machado et al., data mesh is proposed to overcome the limitations of monolithic data architectures. To better understand how data mesh tries to outperform the more traditional data architectures, this section continues on defining data architectures and explains different approaches.

To define what a data architecture is, the definition provided by DAMA International (2017, p. 104) has been selected. *Data Architecture* is defined as the "*... standard terms and designs for the elements that are important to the organization*" where the design "*... includes the depiction of the business data as such, including the collection, storage, integration, movement, and distribution of data.*" To elaborate on standard terms and designs for the elements, the data architecture encompasses a comprehensive pattern of underlying principles, policies, and procedures that guide the data processes across data systems, applications, and data infrastructure.

DAMA International states that the storage of data within an organisation is carried out in a database. Furthermore, a database could be categorised as either centralised or distributed. Whereas a centralised system operates one database on a single system, a distributed system operates several databases across multiple systems. Shakir et al. (2021) and Tiwana (2014) define the single database approach as a monolithic data architecture. In a monolithic data architecture, data is stored and managed centrally. Two of the popular monolithic data architectures are the data warehouse and data lake (Nambiar & Mundra, 2022). Both could be defined as central storage repositories of data (Giebler et al., 2021; Loshin, 2012). Next to the centralised approaches, DAMA International describes that in a decentralised data architecture, it is possible to access data over a larger number of nodes. These nodes could be classified into two types: autonomous (federated) or non-autonomous (non-federated).

Data Warehouse - Data warehouses emerged around 1980 as technology to enable the integration of data coming from various sources into a common data model (Devlin, 2020). According to DAMA International, data warehouses consist of two primary components, a database and software. The software enables collecting, cleansing, transforming, and storing data from different operational and external sources and requesting the data from the data warehouse (Farnum et al., 2019). Martins et al. (2020) state that integrating data from a variety of sources enables the provision of insights into the operational processes and opens possibilities for BI. Dehghani mentions that providing a consolidated point of access to all organisational data is seen as the main benefit of the data warehouse. However, the requirement to modify the data during the ingestion process is a limitation, since the workload for a central data team is heavy to clean all the data for the various organisational departments (Aissi et al., 2022). Moreover, the ingestion process of extracting, transforming, and loading (ETL) causes large complexity (Fahmi et al., 2022). To illustrate, Fahmi et al. mention that organisational operations are supported by various applications, each with its own characteristics. Therefore, different departments within the same organisation end up using different systems depending on their application requirements. The ETL-processes are periodically executed on the operational data sources to extract, transform, and load the data into the data warehouse. Subsequently, various types of analytics are then performed on top of these transformation processes. This shows that managing these ETL-processes for the complete organisation could become challenging (Hechler et al., 2023; Karpathiotakis et al., 2017).

Figure 2.2 presents an overview of the data warehouse architecture (Dehghani, 2022b).

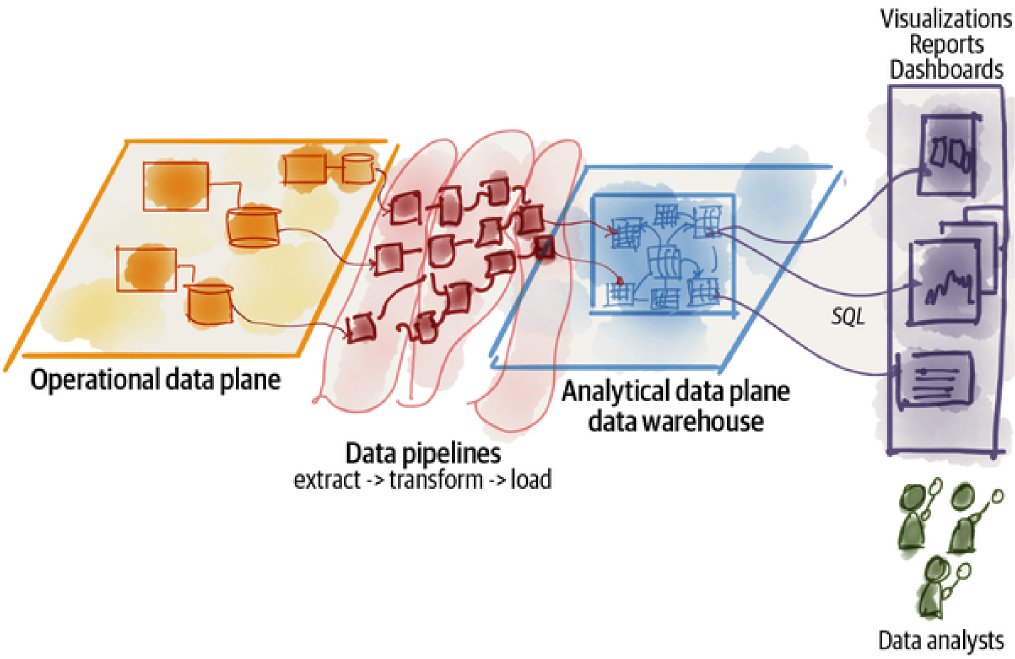


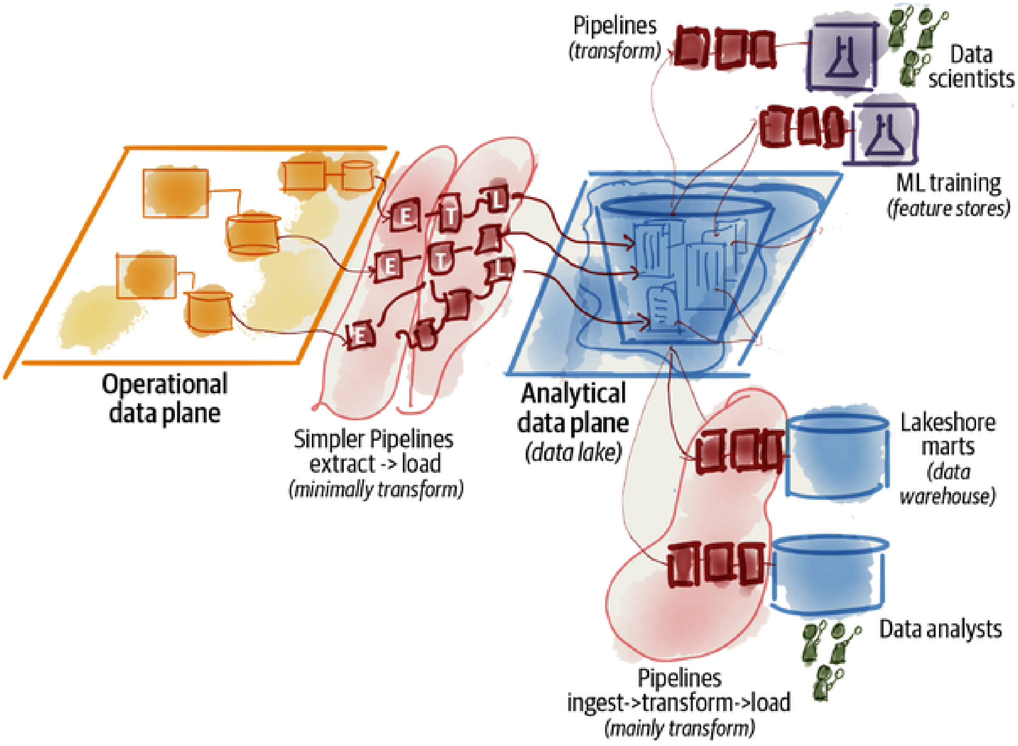
Figure 2.2: Data Warehouse Architecture

The process of data extraction from the many sources towards BI will be described in five steps. First of all, the figure presents in the operational data plane the process where data is extracted from the operational sources. Secondly, the ETL-process is performed in data pipelines. Data pipelines enable the flow of data from the sources towards the data storage. Thirdly, the analytical data plane consists of the data warehouse where data is loaded into warehouse tables. Fourthly, by means of querying, data from the data warehouse could be requested. Lastly, the requested data serves data analysts for reporting, creating visualisations, and designing dashboards.

Data Lake - Aissi et al. state that data architectures continued to evolve to deal with the increased 3 V's of data. Due to the explosion of unstructured and semi-structured data, the data lake architecture was introduced around 2010 (Dixon, 2010). The data lake is an extension of the data warehouse architecture which meets the need to extract insights from big data (Sinan et al., 2022). Giebler et al. and Loshin mentioned that in a data lake, similar as the data warehouse, data is extracted into a central repository. However, the ETL-process has become simpler compared to the data warehouse architecture (Simitsis et al., 2023). Due to data scientists needing access to raw data for optimal ML-model training, extensive up-front data modifications are limited (Khine & Yang, 2018; Wieder & Nolte, 2022). Extensive ETL-processes would lead to slower iterations of ML-model training. However, as the data is present in the data lake, Dehghani mentions that the data architecture still gets extended with additional pipelines to store data at the edge of the lake in lakeshore marts. These marts contain a copy of the stored data from the data lake with a specialised focus on the requests from an organisational department (Belov et al., 2021). As a result, the amount of data to deal with during the training of the ML-models is reduced. Dehghani states that requesting data from these marts therefore allows for a faster data experimentation process instead of extracting data from the central lake. However, the problem is that the pipeline complexity remains, which implies that data still needs to be managed from various sources towards the data lake and from the data lake towards the marts. In addition, the possibility of having multiple data copies across the various marts is seen as a limitation. Data duplication adversely impacts data quality and ML-model training (Choi et al., 2009; Kołcz et al., 2003). Hechler et al. explain that it is an operational challenge to keep data copies consistent over time and that duplicated data is associated with additional data infrastructure and storage costs. Inconsistency in data implies a loss in data quality, which negatively impacts the trust in data.

The process from data extraction towards using data for ML-training, as presented in Figure 2.3, could be summarised in four steps (Dehghani, 2022e). First of all, in the operational data plane is data extracted from multiple operational sources. Secondly, this data is extracted, minimally transformed and loaded by simpler data pipelines compared to the data warehouse architecture. Thirdly, the analytical data plane consists of the data lake where data is stored centrally. Fourthly, data stored in the data lake is used and transformed in data pipelines by data scientists for analytical purposes and ML-model training. Additional data pipelines are created to transform data from the data lake towards the lakeshore marts. These lakeshore marts store data which are usually oriented to specific organisational departments and are used by their applications.

According to Couto & Ruiz (2022) and Dehghani, the data lake architecture creates complex and unwieldy pipelines which deteriorate over time. The complexity arises from the separation of the operational and analytical panels, where data pipelines form the bridge to transform operational data for analytical purposes. In addition, due to data management activities being conducted by a central IT or Analytics team, these pipelines and datasets are becoming unmanaged as the number of sources and the extent to which data consumers want to experiment with data increase, resulting in unwieldy pipelines which deteriorate over time. Consequently, Hechler et al. and Wang & Strong (1996) state that unmanaged data leads to untrusted and inaccessible data throughout the organisation.



From *Data Mesh* (Chapter 8), 2022, O'Reilly. Copyright 2022 by Zhamak Dehghani.

Figure 2.3: Data Lake Architecture

Having defined the data warehouse and data lake as monolithic data architectures, the approaches will be evaluated to assess where these are lacking. According to Dehghani, monolithic data architectures are appropriate as starting point for organisations which need centralised data management. Giebler et al., Loshin, and Martins et al. mentioned that by consolidating data into a centralised repository, a consistent overview of all organisational data could be provided to support creating reports, visualisations, and advanced analytics. However, as the number of data solutions within an organisation increases, it starts showing its limitations. Three main reasons are described by Dehghani for the friction in the monolithic data architectures. First of all, ubiquitous data and source proliferation result in the inability to ingest and harmonise all the data by a central IT or Analytics team. Alrehamy & Walker (2018) explain that the manual tasks for a central team pose a huge burden over time. Moreover, since the data activities are performed outside the domains, there is also a disconnect between the people who understand the data and the people who actually process the data. The centralised ownership and responsibility create the bottleneck as more manual work and coordination are needed. Secondly, increasing needs within an organisation to experiment with data and the use case proliferation result in an ever-growing number of data transformation processes. As a result, the pattern of ETL-processes is becoming highly complex. Thirdly, monolithic data architectures cannot easily scale. These gradually become slow, expensive, and hard to maintain. Alrehamy & Walker explain that the scalability challenge arises as soon as the number and variety of data sources increases. As a result, Dehghani argues that the monolithic data architecture becomes the bottleneck for organisations in case of continuous change of the data environment.

Dehghani explains that the monolithic approaches with a high level of pipeline-based integration remain fragile and hard to maintain in the nowadays highly volatile environment while data experimentation is becoming more important than ever before. The monolithic architectural and organisational structure does not deliver the value needed for organisations to become data-driven. The need for a data architecture that is able to respond to complexity and scale, and meet the aspiration of data experimentation asks for change.

Data Mesh - As described by DAMA International, databases were classified as either centralised or distributed. Due to the limitations of the centralised data warehouse and data lake, a decentralised approach to alternatively manage data is introduced. Dehghani (2022a, chap. 1, para. 1) introduced data mesh as "*... a decentralized sociotechnical approach to share, access, and manage analytical data in complex and large-scale environments — within or across organizations*" and as "*... a new approach in sourcing, managing, and accessing data for analytical use cases at scale.*" The ultimate objective of data mesh is to "*... increasing the ability of organizations to utilize data for analytical purposes and get value from their data at scale, aligned with organizational growth and complexity.*" Since data mesh is neither a data architecture, a list of principles, nor an operating model, Dehghani classified data mesh as a socio-technical paradigm.

The data mesh paradigm is the convergence of four main dimensions, which are *domain oriented decentralised data ownership and architecture*, *data as a product*, *self-serve data infrastructure as a platform*, and *federated computational governance*. These dimensions drive the organisational structure and technical architecture to increase the value of data at scale, sustain agility during organisational growth, and embrace change in a complex data environment.

Domain Oriented Decentralised Data Ownership and Architecture - The ownership of data needs to be decentralised to the domains closest to the data. In other words, data responsibilities such as data collection, transformation, integration, quality assurance, security, and sharing lie with those who are most familiar with the data. As a result, the domain-oriented data will be managed decentrally and independently from a centralised IT or Analytics team.

Data as a Product - Data provided by the domains need to be treated as a product. This means that product thinking will be applied to how data is modelled and shared. In addition, data products need to meet all usability attributes to guarantee data products are uniquely valuable. Data as a product adheres to be discoverable, addressable, understandable, trustworthy, natively accessible, interoperable, and secure.

Self-Serve Data Infrastructure as a Platform - The data infrastructure integrates the operational and analytical capabilities into a self-serve data platform. The platform empowers the domain's cross-functional teams, with decentralised technologies, to build and share interoperable data products autonomously to serve domain-agnostic use cases.

Federated Computational Governance - Federated computational governance refers to the decision-making model that balances autonomy, agility, and local decision-making power of domains while creating and adhering to defined global rules. In addition, the governance policies will be automated. In other words, the governance operating model, which impacts how an organisation operates around its data and establishes an organisational governance structure (Brennan et al., 2018), relies on automation by computational policies. Rahimzadeh et al. (2022) explain the automation of the governance operating model as a series of pre-programmed steps that will be followed by automated processes to ensure compliance with the data governance policies. This implies that computational policies will be established for data products, via the platform services, to assure data is secure, compliant, of quality, and usable.

Figure 2.4 provides an overview of the four data mesh dimensions (Dehghani, 2022d).

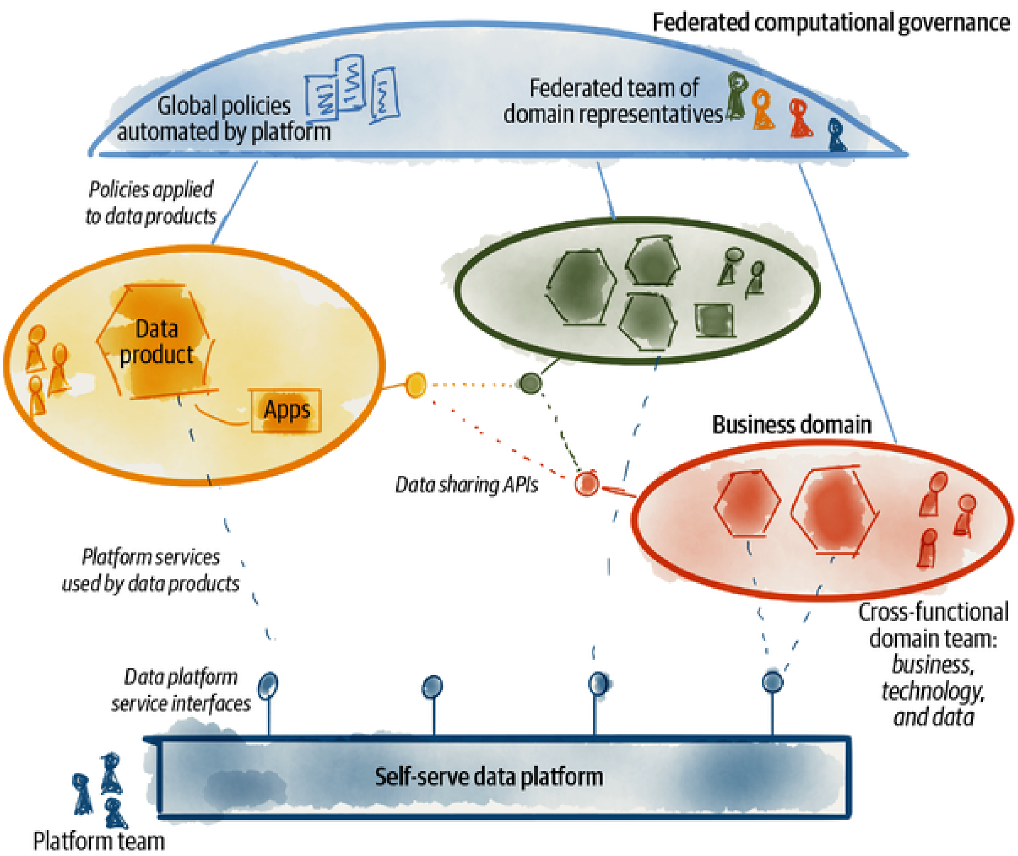


Figure 2.4: Data Mesh Architecture

The dimension of *domain oriented decentralised data ownership and architecture* is visualised by the yellow, green, and red ovals which represent the domains in an organisation. The domains take ownership of their *data as a product*. The data architecture is aligned with the decentralised domain structure. The bottleneck of a centralised IT or Analytics team and the disconnect between the people who understand the data and who transform the data are therefore removed. As the domains are responsible for performing their data activities, cross-functional teams need to be established. Cross-functional teams need to include business, technology, and data professionals. Furthermore, the domains are connected via data-sharing application programming interfaces (API's), which are part of the *self-serve data infrastructure as a platform*. The platform is centrally provided by a platform team, which enables the domains to autonomously share data products. The data products are owned by data product owners who take the ownership of data. The data product encapsulates the ETL-process, which means it transforms operational data into analytical data through an internal pipeline. This process is internally managed by the domain teams, eliminating the requirement for centrally managed pipelines between the operational and analytical planes. This aim arises from the desire to have analytical data reflecting as best as possible the activities within the organisation, ensuring a close link between analytical and operational data. At last, the *federated computational governance* balances global authority by a federated team of domain representatives and federated decision-making initiated at the domain level. Mesh-wide standardisation is obtained by having global communication protocols in place. These protocols govern how data products express their semantics, format, query language, and what service levels objectives (SLO) each guarantees to enable data product interoperability. Global policies need to be automated via the platform and in the data products to reduce manual processes.

Dehghani argues that all dimensions are collectively necessary, complementary, and dependent on each other for the proper functioning of a data mesh. To illustrate this with an example, *domain oriented decentralised data ownership* could lead to data siloing within domains, resulting in duplicated work and increased costs of data ownership. Data siloing could be prevented by the principle of *data as a product*, which demands that domains take the responsibility to share their data with other domains possessing product-like qualities, such as understandability and trustworthiness. In addition, the *self-service data infrastructure* provides the platform for domain teams to develop, share, and use data products from one another. By assigning the responsibility and accountability of data ownership to the domains, they are also entrusted with the task of fulfilling that responsibility. Dehghani and Grossman (2023) refer to a positive network effect that could arise when all domains collaborate, with connectivity between domains exchanging data products. The larger the network and the more connections established, the more data could be shared between domains, from which value could be derived for the organisation. However, if domains fail to fulfil their responsibility of creating and providing quality data products to other domains, a situation may arise where the exchange of data products stagnates, thus eliminating the network effect, or even resulting in unusable data products. This would lead to the fact that poorly managed data has disadvantages, such as data duplication and inconsistencies, which in turn have a negative impact on data quality, mention Choi et al. As a result, Hechler et al. and Kołcz et al. argue that ML-models will be trained less effectively, and costs will increase for data infrastructure and storage. Consequently, Wang & Strong state that trust and accessibility of data within the organisation decline. In short, this demonstrates the importance of collaboration between domain teams, as well as the presence and functioning of the four dimensions in organisations that are interconnected, complementary, and dependent on each other.

2.1.4. Defining the Paradigm Shift

Data mesh introduces multidimensional organisational and technical shifts with respect to monolithic data architectures. The transition towards a data mesh is therefore classified as a paradigm shift. Six main transitions are identified by Dehghani (2022f), which are visualised in Figure 2.5.

Organisationally - Whereas in a monolithic data architecture, the ownership is managed by a central team of data specialists, data mesh pushes the ownership and accountability back to the domains which produce or use the data. The responsibilities are therefore allocated decentrally.

Architecturally - Data collection shifts from the monolithic data warehouse or data lake towards a distributed mesh of data products. The architecture matches the decentralised domain-oriented structure.

Technically - Data as a byproduct of code shifts towards data products in which the data and code are provided as one autonomous unit. More specifically, data mesh conceptualises each data product as an architectural quantum unit, wherein it holds the domain-specific data alongside the code responsible for the data transformations and shares the metadata and associated governance policies. In other words, everything that is relevant to fulfilling the usability attributes must be included in the data product.

Operationally - Global top-down governance shifts to federated computational governance, which balances global policies with domain-level authority. Computational refers to the need to integrate governance policies in the data infrastructure and data products to enable automation.

Principally - Product thinking enables the change from data as an asset to be collected towards data as a product to be shared and connected. Data mesh encourages sharing data products to serve domain-agnostic use cases.

Infrastructurally - The structure of a fragmented operational and analytical data plane is replaced by an integrated self-serve platform.

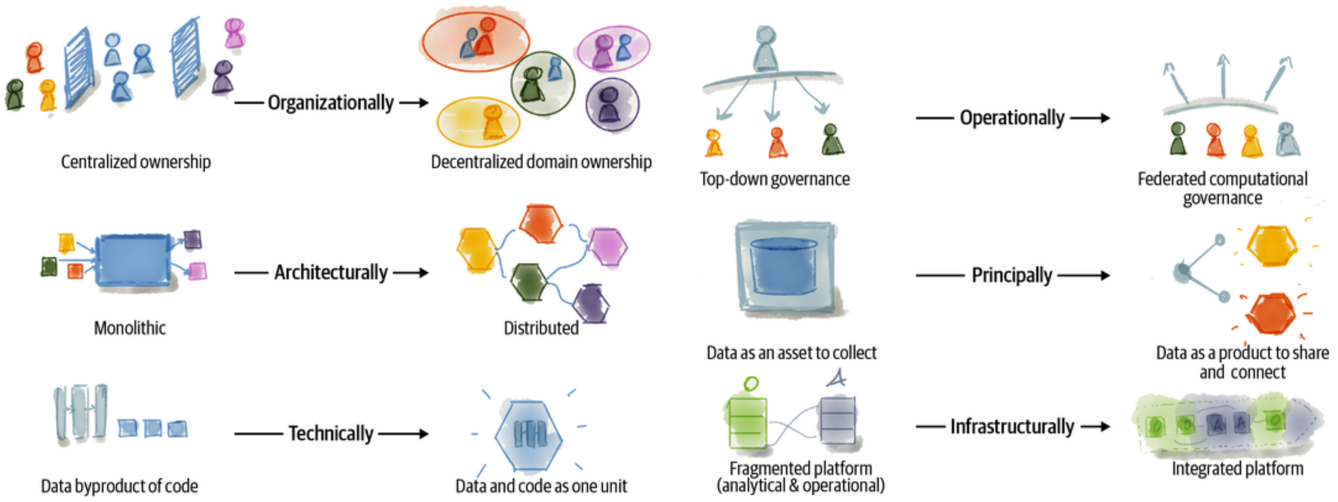


Figure 2.5: Data Mesh Dimensions of Change

While the shifts are explained, further explanations will be provided about the shifts *operationally* and *principally*, regarding respectively the functioning of a balance between central and decentralised policies, and the potential impact of a shift in data ownership when considering data as an element of power and accountability.

According to shifting operationally, the governance operating model of the data mesh consists of three complementary pillars states Dehghani. Firstly, systems thinking is required, where the mesh is seen as an ecosystem of interconnected data products and platform systems, along with their independent yet interconnected domain teams. The aim is to identify leverage points and feedback loops to control the behaviour of the mesh in order to maintain an equilibrium between global and domain autonomy. For instance, establishing security and legal standards centrally would be desirable since these regulatory topics encompass the entire organisation (Khatri & Brown, 2010). On the other hand, Dehghani mentions that domains are granted autonomy and responsibility for the majority of policies within their sphere of influence and control. For example, domains could appoint data product owners themselves or decide on the frequency of updates of their data. These examples show that decision-making is situated as close as possible to the individuals impacted by these decisions. Secondly, a team comprising domain product owners, platform representatives, and SME's, such as legal and security professionals, needs to be formed to centrally establish policies. It is important to note that the approach to balancing central and decentralised policies is influenced by various factors, such as organisational culture and values (Gupta, 2020; Hendriks, 2023). To illustrate, in the case of a hierarchical organisational culture, there may be a tendency to favour centralised governance, whereas, in a cooperative organisational culture, responsibilities are more likely to be distributed among different domains (Otto, 2011; Weber et al., 2009). Thirdly, from a practical and implementation perspective, data mesh governance relies on embedding governance policies in each data product in an automated and computational manner. To illustrate, data quality checks could be automated, ensuring these are performed after each data update (Borisyak et al., 2017). Additionally, user authentication functionalities could be integrated for secure data accessibility (Johri et al., 2017).

According to the shift principally, the transition from considering data as an asset to be owned by a central IT or Analytics team, to viewing it as a product managed by domain product owners, also represents a shift in data ownership as an element of power and accountability (Asswad & Gómez, 2021; Certybox Education, 2023; Krystek et al., 2023). For an organisation, this shift could be desirable and reasonable as it aims to promote data product exchange across domains to obtain network effects, where data products are owned decentrally and managed by those most familiar with the data, as stated by Dehghani, Grossman, and Wider et al. (2023). However, the ease with which this shift towards data as a product to be shared could be made varies across organisations, where it may not be suitable at all for some organisations (Hokkanen, 2021). To illustrate with an example from Asswad & Gómez, an organisation that handles a significant amount of sensitive information, such as in the healthcare sector, is referred to as having numerous privacy issues related to changing data ownership. Strict laws on data usage, accessibility, and privacy could make it difficult to transfer ownership of the data.

2.2. Outcomes of Data Mesh

This section will explain the outcomes of data mesh, including the benefits presented in Section 2.2.1, the challenges related to data mesh, and changing a data architecture presented in Section 2.2.2 and 2.2.3 respectively.

2.2.1. Benefits of Data Mesh

Data mesh is presented as the solution for organisations that experience complexity and scale, where monolithic data architectures have become the bottlenecks in their ability to extract value from data. Dehghani assumes that data mesh achieves the following three main outcomes: responding gracefully to change in a complex organisation, sustaining agility in times of organisational growth, and increasing the value out of data relative to its effort and investments. Before presenting the outcomes, it is important to reiterate the statements from Driessen et al. (2023) and Machado et al., referring to there is still a long way to go when it comes to concluding the functionality of data mesh. Moreover, Goedegebuure et al. (2023) emphasised the current lack of academic publications about data mesh. Section 1.2.1 also showed the relatively low number of scientific contributions related to data mesh. In addition, data mesh also lacks empirical insights from the field (Bode et al., 2023; Butte & Butte, 2022). Altogether, when interpreting the expected outcomes as presented by Dehghani, it is important to take into account the novelty of the approach and the limited academic and empirical research available on these outcomes.

Respond Gracefully to Change in a Complex Business - Data mesh embraces change despite increased organisational complexity. Responding gracefully to change will be enabled by four reasons. Firstly, managing complexity becomes easier when technology and business professionals are aligned with the analytical data. The domains control and manage their operational data and applications, supported by technology professionals. Secondly, data mesh closes the gap between analytical and operational data. This desire has arisen because analytical data needs to reflect as best as possible what happens within the organisation. For this reason, analytical data must be as close as possible connected to operational data. Data products will therefore embed and abstract the data transformation as an internal pipeline managed by the domain teams. Thirdly, domains are empowered to model their data without the help of a central IT or Analytics team. This means that data changes are localised to the domains. Fourthly, accidental complexities, in terms of pipelines and duplicated data will be reduced, since data products will be natively accessible. Data will be shared without the need for intermediary pipelines. The original form of data will be maintained and different copy versions will be avoided.

Sustain Agility in the Face of Growth - Data mesh reduces bottlenecks, coordination, and synchronisation issues. Pushing back the data responsibilities towards the domains achieves agility outcomes. In the case of acquisitions, new organisational domains, new products, or international expansions, the organisational data structure will be able to adapt easily. Sustaining agility in the face of growth is enabled by four reasons. First of all, the centralised IT or Analytics team limits agility as the number of sources and data use cases increases. From a technical perspective, data mesh enables domains to directly discover and share data products from various sources by themselves. This means that domains operate autonomously with minimal dependencies. Secondly, data mesh aims to reduce architectural coordination. The process of ingestion, processing, and serving in monolithic data architectures asks for high levels of coordination between functional teams and central data specialists. Moving away from the technical portioning to domain-oriented partitioning removes the friction of coordination. Thirdly, manual governance is seen as a major bottleneck in monolithic data architectures. By embedding the policies as code for automation in the data infrastructure and each data product, the manual inefficiencies will be reduced. Fourthly, data mesh balances the freedom in decision-making by domains and central autonomy by global governance. This balance would lead to better individual team performance while preventing team isolation and duplicated efforts.

Increase Ratio of Value from Data to Investment - Data mesh aims to increase the ratio of value from data to investment. This will be accomplished in three different ways. Firstly, data mesh aims to reduce the complexity of the current data management technologies. Open and standardised interfaces are proposed as a solution to create a more collaborative ecosystem of technologies. Secondly, data product thinking would help increase the value of data. Data product thinking means shifting the focus from data as an asset to be collected towards treating data as a product to share. This shift would reduce the effort and costs due to network effects. Thirdly, more value will be delivered by data mesh as data is connected beyond its organisational boundaries. Data mesh provides a set of interfaces that allow anyone to access data products regardless of their physical location.

2.2.2. Challenges of Data Mesh

Next to the benefits, there are several challenges presented in literature according to implementing data mesh. First of all, Podlesny et al. (2022) discussed potential issues that emerge in guaranteeing privacy across the different domains. Linking data products from different domains could be exploited to subvert privacy. Bode et al. agree on this concern by stating that mainly security, regulatory, and compliance issues will occur when data will be organised decentrally. Employees in the domains are unaware of what data is protected and regulated, which could result in non-compliance. Vestues et al. (2022) provide the General Data Protection Regulation (GDPR) as such regulation that may problematise the data sharing mindset. Secondly, Joshi et al. (2021) mentioned the challenge of successfully attaining intelligent federated governance. Finding a balance between global agreement and federated decision-making in a dynamic organisational environment is a moving target. Thirdly, Vestues et al. state that data mesh requires extra work by the domains and all domains have a different set of data skills. As a result, this could lead to the risk of erroneous data and reduced data quality. Vestues et al. doubt about the competence of domain teams to create data products on their own. It is unclear which skill set is required, what the costs are, and how cross-functional teams will be composed. Fourthly, Bode et al. and Vestues et al. expect a lack of incentives to build data products which will be consumed by others. If domains do not automatically receive compensation for data provision efforts to serve other domains, to what extent are domains willing to put effort into making the data product useful and understandable for others? At last, introducing data mesh requires a new mindset and terminology. Bode et al. expect misinterpretation due to a lack of common understanding.

2.2.3. Challenges of the Paradigm Shift

Next to the data mesh challenges, initiating and changing a data architecture approach in general also exposes challenges. From a financial, technical, and human perspective, Bode et al. state that resources are needed to meet expectations and enable change. According to change, also acceptance issues and resistance could be expected. In addition, DAMA International mentions that reorganisation asks for long-term management support. At the same time, it is needed that managers are experienced with the new approach and are able to convince others about the potential of data mesh. Lastly, data mesh requires a cultural shift, which means employees have to change their behaviour and mindset. which is often perceived as a difficult process.

Phase B

Objective for a Solution

3

Maturity Assessment Models

Chapter 1 mentioned the current lack of guidance on how to transform towards data mesh. The *Objective for a Solution* phase B infers the desirability for this guidance. It is expected that guidance, in the form of a maturity assessment model, would help organisations implement data mesh. This chapter will therefore provide more explanation about the contribution of maturity assessment models and will discuss their applicability to data mesh. The structure of this chapter follows three sections. Section 3.1 will define what is considered by a maturity assessment model to obtain a common understanding. To evaluate the contribution of maturity assessment models to organisations, the goals and drivers will be examined in Section 3.2. Existing big data, data management, and data analytics maturity assessment models will be discussed in Section 3.3 to learn about the best practices regarding the main concepts and features. At last, this chapter closes by concluding how the DMMAM contributes to the data mesh implementation process. In addition, it will be concluded which design elements are important to take into consideration while developing the DMMAM. The sub-question, sub-objective, and sub-deliverable of this chapter are provided below.

Sub-Question 1

How does a data mesh maturity assessment model contribute to the data mesh implementation process?

Sub-Objective 1

Explore how the data mesh maturity assessment model helps organisations implement data mesh.

Sub-Deliverable 1

Relevance of a data mesh maturity assessment model to the implementation process.

3.1. Defining Maturity Assessment Models

To obtain a common understanding of what is considered a maturity assessment model, this section will explain its definition in Section 3.1.1, the model elements in Section 3.1.2, and the assessment activities in Section 3.1.3.

3.1.1. Definition

The concept of the Capability Maturity Model (CMM) is coined by the Software Engineering Institute in 1984 to provide a methodology for improving the software development process (Kitson & Masters, 1992). Since the success of the CMM, there has been a surge in the development of maturity models (Adekunle et al., 2022). Becker et al. (2009) state that over a hundred maturity models have been developed in recent years in the field of information management. This number indicates the increased need for these models (Steenbergen et al., 2010).

To define maturity assessment models, the definition presented by DAMA International (2017) will be used since their concepts are considered as the standards. DAMA International defines maturity assessment as an approach to process improvement, based on a model that describes how characteristics evolve over levels, which indicate an organisation's current capabilities and the desirable states. In other words, Mettler et al. (2010) and Schumacher et al. (2016) explain that all characteristics related to a topic will be evaluated for an organisation to obtain an understanding of its current and target level of maturity. DAMA International states that progression over the levels happens in a set order, which implies that it is not possible to skip a level. To elaborate, for each consecutive maturity level, the characteristic's process becomes more developed, thereby demonstrating a logical progression through stages (De Bruin et al., 2005). Król & Zdonek (2020) define this process as the path to perfection. In short, a maturity assessment helps an organisation to provide insights into its current capabilities and the desirable states by assessing its maturity levels to open opportunities for improvement (Tarhan et al., 2016).

Pöppelbuß & Röglinger (2011) describe that the maturity assessment model could be used as *descriptive*, *prescriptive*, and as *comparative* model. Whereas the descriptive model assesses the *as-is state*, the prescriptive model is responsible for the connection between organisational performance and enhancement of maturity, aiming for reaching the *to-be state*. The maturity assessment model considered as comparative model enables capability benchmarking with industry competitors.

3.1.2. Model Elements

According to Al-Sai et al. (2022), DAMA International, Korsten et al. (2022), and Lasrado et al. (2015), essential elements of a maturity assessment model are the *dimensions*, represented by *characteristics*, and evaluated across different *maturity levels*. In addition, *criteria and requirements* are provided for all characteristics over each maturity level. Al-Sai et al. explain that the dimensions are the main principles of the topic being assessed, which are further elaborated through a greater number of characteristics. Subsequently, these characteristics will be evaluated during the assessment along the maturity scale. Becker et al. and Mettler et al. mention that the degree of maturity defines the state of progression within a fixed scale, from an initial point to an endpoint, where the initial point corresponds to the lowest maturity level and the endpoint represents the highest maturity level. DAMA International states that maturity assessment models usually provide approximately five levels, each with its own criteria and requirements that cover the development process. However, De Bruin et al., and Proença (2016) state that the number of maturity levels could differ, depending on the development motivation. Most importantly, De Bruin et al. mention that the maturity levels need to be distinct, well-defined, and provide a logical progression through stages. Furthermore, it is worth noting that existing maturity assessment models, including those designed for data lakes and data warehouses, do not establish a fixed initial point, such as either level 0 or level 1 (Fahmi et al., 2022; Halper & Stodder, 2014; Miller et al., 2011; Sitarska-Buba & Zygała, 2020). The emphasis lies more on the level of meaning and definition rather than their numerical value. At last, the criteria and requirements refer to an activity, tool, standard, and people or resource mentions DAMA International. Activities are evaluated according to the degree the activity is defined, executed, and performed. In the case of a tool, it could be evaluated based on its availability, the extent to which it provides automation, and whether it delivers effective and efficient outcomes. Standards are often initiated to support an activity. Maturity levels could refer to the initiation of standards, how well the standards are documented, the extent the standards are automated, and whether monitoring is in place to safeguard compliance. People or resource refer to the extent to which specific skills, training, and knowledge are available, or roles and responsibilities are defined.

However, considering that Dehghani (2022a) presents data mesh as a fundamentally new approach for organisations to manage data, it is important to determine in advance the extent to which it would be acceptable to adopt these usual practices. To argue that these model elements are also applicable in the context of data mesh, findings from De Bruin et al., García-Mireles et al. (2012), Lahrmann & Marx (2010), and Wendler (2012) will be explained. De Bruin et al. made a large effort to generalise the phases of developing a maturity assessment model from various fields. According to dimensions and characteristics, De Bruin et al. mentioned that the maturity assessment model for any domain of interest could be represented by its domain components and sub-components, enabling organisations to achieve a more detailed understanding of the concepts involved. According to maturity levels, De Bruin et al. continued that the representation of maturity levels as a progression of one-dimensional linear stages is widely adopted, serving as a foundation in many established models. According to criteria and requirements, García-Mireles et al. and Wendler, who conducted systematic literature reviews (SLR's) on the development of maturity assessment models across various fields, further emphasise that it is common practice to add descriptions of how each characteristic evolve across the different maturity levels. Lahrmann & Marx, who provided methodical guidance for maturity assessment model evolution and extensions, mentioned that the structure of existing maturity models could be transferred toward new areas of interest. In conclusion, regardless of data mesh being a fundamentally new approach, it is customary to include these elements in any maturity assessment model. García-Mireles et al., Pino et al. (2008), and Staples & Niazi (2008) argue that this commonality is the result of that all these models are based on common standards for developing maturity assessment models, such as *CMM*, *ISO/IEC 15504*, or *CMMI-DEV* (CMMI Institute, 2012; ISO, 2012; Paulk et al., 1993). Since Steenbergen et al. highlighted that the proliferation of developed maturity models is driven by the demand for such models, these standards will be adhered to during the development of the DMMAM.

3.1.3. Assessment Activities

DAMA International mentions that performing an assessment for an organisation requires planning. The total assessment consists of five consecutive steps, namely planning the assessment activities, performing the maturity assessment, interpreting the results, creating a targeted programme for improvements, and conducting re-assessments.

I. Planning Assessment Activities - Planning the assessment refers to setting up the assessment objectives for the organisation, selecting an appropriate maturity framework, defining the assessment scope, defining an interaction approach, and planning the communications.

II. Performing Maturity Assessment - According to De Bruin et al., there are three methods of conducting an assessment: self-assessment, certified professional-assisted, or third-party-assisted. In the case of a professional-assisted assessment, it involves the assistance of a certified organisation, such as the model developing organisation. On the other hand, third-party assistance could be provided by an external consultancy firm. While conducting the assessment, inputs could be obtained through individual interviews or focus groups. Al-Sai et al., García-Mireles et al. (2012), and Steenbergen et al. state that the application could also be supported by an assessment tool, questionnaire, or checklist. The scope of the research determines the level of detail of the assessment and the number of respondents needed across various departments. Furthermore, DAMA International mentions that maturity ratings across the participants are often slightly different. Next to interviewing sufficient respondents, a discussion and rationalisation session is beneficial to reconcile the ratings. In the end, it is the aim to find a consensus between the assessor and the participants regarding the maturity states in case of dissension.

III. Interpreting Results - Interpretation of the results refers to observing the as-is state and identifying opportunities for improvement. Next to assessing the current maturity states, asking about target maturity states for a given scope is beneficial to identify the maturity gaps. These maturity gaps will be the starting point for creating roadmaps and identifying what resources are needed to close these gaps. For an organisation, it is interesting to obtain insights into the drivers for doing the assessment, the overall results, the results per dimension or characteristic, the maturity gaps, and recommendations to close the gaps.

IV. Creating a Targeted Programme for Improvements - Recommendations from the assessment need to be actionable. Actions need to be identified from which targeted programmes could be created. These roadmaps will provide targets and the tempo for change.

V. Re-assessing Maturity - To guarantee a cycle of continuous improvement, re-assessments need to be conducted at regular intervals. Stoiber et al. (2023) concur with DAMA International that a single maturity assessment of capabilities fails to capture the ongoing evolution of capabilities within dynamic organisational environments. In addition, measuring progress maintains commitment and enthusiasm across the organisation. But, when the improvement process is lacking, this will also be noticed by the re-assessments.

3.2. Goals and Drivers to Use Maturity Assessment Models

This section will look into the goals and drivers to use maturity assessment models to evaluate their contribution to organisations. Important to mention beforehand that Section 3.1.2 addressed the question of whether it would be acceptable to apply the design elements found in existing maturity assessment models to the development of the DMMAM. It has been determined that the maturity assessment model structure is not dependent on its respective field of application. Wendler, who explored the benefits and purposes of maturity assessment models across twenty domains, asserts that also benefits and purposes are not restricted to any specific domain.

Belghith et al. (2021), Mettler et al., Schumacher et al., and Tarhan et al. explained that the main objective of performing a maturity assessment refers to the aim to evaluate the current state of capabilities in order to identify, prioritise, and implement improvement opportunities. Kerzner (2001), Król & Zdonek and Lemsa (2021) add to this aim that the strengths and weaknesses of an organisation will be assessed. Al-Sai et al. and Helfert & Donnellan (2012) state that the model functions as the scale for evaluating the current state on the path of transformation. Altogether, the three main goals initiated by DAMA International will be considered as the main goals to conduct a maturity assessment. Comprehensively discovering and evaluating the activities of interest across an organisation is the first goal. By performing a maturity assessment, an organisation will be able to identify its as-is state. The second goal refers to educating stakeholders about practices, concepts, and principles as well as identifying roles and responsibilities. Establishing or enhancing the organisational programme in support of strategic and operational goals is considered the third goal. The programme represents a roadmap to guide towards the desirable maturity states. The assessment is a starting point for the development of plans, actions, strategies, and roadmaps to reach the to-be state.

Next to the goals of performing a maturity assessment, several main drivers for organisations to conduct a maturity assessment are provided in the literature. First of all, DAMA International mentions that *regulation oversight* could require minimal maturity levels for specific data management characteristics. Secondly, Al-Sai et al. and DAMA International state that organisations want to assess their *readiness* for improving their processes. Recognising the need to improve the practices starts by assessing current maturity states. Thirdly, data management challenges could occur in times of organisational change, such as a merger. To meet these challenges, DAMA International mentions that an assessment provides relevant *input for planning*. These challenges could be addressed by assessing the current state to make better decisions while going through a transition. Fourthly, Al-Sai et al. motivate that the maturity model could be useful to *track organisational performance*. By means of the model, Belghith et al. explain that the ability to achieve pre-determined goals could be assessed. For example, by having implemented a new technology, an organisation wants to understand how this advancement will lead to successful adoption. Fifthly, Al-Sai et al. mention that the maturity assessment could also be considered a *classification tool*. Meaning that the model determines the current maturity and the necessary values for risk, quality, cost, and return on investment (ROI) to attain the desired levels. Sixthly, Kerzner states that maturity models are useful for *benchmarking* an organisation with its industry competitors. McCormack et al. (2009) and Lemsa found that using a benchmark tool helps organisations understand their position and grants a competitive advantage through adequate tracking of their industry peers.

3.3. Benchmarking Data Maturity Assessment Models

According to Becker et al. and Adekunle et al., plenty of maturity assessment models have been designed since the introduction of the CMM. To inform the design process of developing the DMMAM, this section aims to draw insights from the main concepts and features of the many existing maturity assessment models. Evaluating these models as a benchmark will help acquire design knowledge to inform the development of the DMMAM (Dym et al., 2013).

To search for existing maturity assessment models, the Scopus database was used. When searching for existing maturity assessment models, a systematic literature review (SLR) is considered helpful for obtaining a comprehensive overview of available models. Consequently, the search term *"Maturity" AND ("Model" OR "Assessment") AND "Literature Review"* was applied within *Article Title, Abstract, and Keywords*. This resulted in 870 publications. Due to the high number of results and the assumption that it would be more beneficial to obtain an overview of maturity assessment models more closely related to data mesh, the scope was narrowed down to only examining maturity assessment models that specifically pertain to *data*. The revised search term *"Data" AND "Maturity" AND ("Model" OR "Assessment") AND "Literature Review"*, applied within *Article Title, Abstract, and Keywords*, resulted in 255 publications. Since it was aimed to obtain an overview of state-of-the-art literature, the search term was limited to only documents from 2020 and beyond. Furthermore, the search only focused on *English* publications within the *Computer Science* subject area. This final search resulted in 74 publications, which were sorted based on *relevance* in Scopus. Subsequently, the abstracts of the first 20 publications were analysed. In the end, the publications from Al-Sai et al., Belghith et al., and Król & Zdonek seemed most relevant to inform the design process for developing the DMMAM. These authors conducted an SLR in which respectively existing *big data*, *data management*, and *data analytics* maturity assessment models were evaluated without any industry focus. The other suggestions were specifically focused on an industry or application. As this research aims for a model which is not specific to any particular organisation or industry, these publications were excluded. The selected SLR's are also expected to be more useful as these cover aspects which are closely related to data mesh. To explain, Strengholt (2023) states that data mesh is an approach to data management aiming for managing data at scale. Dehghani mentions that data analytics and big data are concepts both embraced by data mesh. Dehghani argues that data analytics and big data are significant contributors to the existence of data mesh. To bring it all together, these SLR's will be analysed and compared to obtain an understanding of what data maturity assessment models already exist and presents the commonalities and differences. Appendix A shows the covered models from Al-Sai et al., Belghith et al., and Król & Zdonek in an adapted format, namely by mentioning the *name*, *author*, *number of maturity levels*, and *number of dimensions*. The decision to include these elements was based on the fact that these were also the focus of comparison in the individual SLR's. Consequently, this information was available, allowing for consistent and uniform identification of these elements. The purpose of including these elements is to understand the common elements and how the DMMAM-design relates to existing models in this field. Furthermore, Section 3.3.1, 3.3.2, and 3.3.3 will explain the set of maturity assessment models, best practices, and limitations in the context of big data, data management, and data analytics respectively, as it will provide lessons for the design process.

3.3.1. Big Data

Al-Sai et al. performed an SLR aiming for covering existing *Big Data Maturity Assessment Models* (BDMAM's) and addressing the model limitations. In total, 15 BDMAM's were identified, compared, and discussed, which are presented in Table A.1. The main findings will be mentioned. Firstly, the existence of 15 BDMAM's is considered a significant quantity, especially due to their exclusive focus on big data. Secondly, the list of authors who developed the maturity assessment models encompasses numerous individual academic researchers. Thirdly, the number of maturity levels ranges from 4, 5, to 6. Lastly, on average, the models consist of 5 dimensions. Next to the model elements, 12 limitations of the BDMAM's were presented by Al-Sai et al. Table A.2 shows and ranks the limitations based on their occurrence. The most important lessons to be learnt refer to enabling the assessment as self-assessment, supporting the application through a digital questionnaire that includes formulated control assessment questions for the identified characteristics, explaining the development procedure, identifying dimensions and characteristics, providing documentation about the model, describing the assessment method, presenting a visualisation report with the outcomes, and validating the model by experts and through assessments conducted at organisations. To highlight, the lack of validation in the majority of developed maturity assessment models is regarded as a significant limitation state Santos-Neto & Costa (2019), Tarhan et al., and Wendler.

3.3.2. Data Management

Belghith et al. conducted an SLR in which maturity assessment models either for or related to data management were reviewed. In their paper, a comparative analysis was performed to address the main concepts and features associated with these models. The analysis compared 22 *Data Management Maturity Models* (DMMM's) to examine the commonalities and differences. Subsequently, these 22 models were subdivided into 6 families, which are *Data Management*, *Data/Information Governance*, *Software Development*, *Digital Assessment*, *Analytics*, and *Business Performance*. Table A.4 presents the 22 models examined by Belghith et al., from which several findings will be mentioned. Firstly, the existence of 22 DMMM's is considered a significant number, especially considering their exclusive focus on data management. Secondly, a majority of these models have been developed by non-profit associations dedicated to advancing data management and analytics knowledge like *DAMA International*, *CMMI Institute*, and *EDM Council*; technology or consultancy firms like *IBM*, *Gartner*, *Deloitte*, and *DELL*; and research institutes such as the *School of Information Studies at Syracuse University*, *Stanford University's Data Governance Office*, and the *Software Engineering Institute of Carnegie Mellon*. Thirdly, almost all maturity assessment models have 5 levels, except the model from *DataFlux Company* and *DAMA International*, which have 4 and 6 maturity levels respectively. Lastly, the number of dimensions covered varies between 3 and 15, with an average of 5.

Table A.5 shows the strengths and weaknesses of the DMMM's, according to the different families, as presented by Belghith et al. The most important findings of this research will be mentioned. The strengths identified encompassed the models' ability to offer guidance and detailed information about their features, provide insights into the current state of maturity and recommendations for its evolution, assist in risk assessment and resource allocation, and enable comparisons of maturity outcomes across different organisations. However, a notable observed weakness was that certain models require extensive resources and knowledge for conducting the assessment.

3.3.3. Data Analytics

Data Analytics Maturity Assessment Models (DAMAM's) were evaluated by Król & Zdonek to review, characterise, and comparatively analyse the model features. Their paper described and analysed in total 11 DAMAM's. Table A.3 presents these models, from which the main findings will be discussed. Firstly, the 11 DAMAM's are considered as an extensive list as these only focus on data analytics. Secondly, all of these models, except for the DAMAM by *Aryng LLC*, consist of 5 maturity levels. The *Analytics Maturity Quotient Framework* developed by *Aryng LLC*, instead of using a discrete scale, employs a continuous maturity scale ranging from 0 to 10. Thirdly, the number of dimensions covered by these DAMAM's is 5 on average. In addition to the findings from Table A.3, Król & Zdonek observed that while comprehensive descriptions of an organisation's maturity levels were available in all models, detailed explanations regarding the criteria and requirements for assigning an organisation to a specific maturity level or the assessment process itself were often lacking. In some instances, these were intentionally omitted, as assessments were conducted on a commercial basis.

By providing the existing data maturity assessment models, an understanding is obtained about which maturity assessment models are currently used in practice and what aspects are perceived as benefits or limitations. The insights relevant to the DMMAM-design will be discussed in the conclusion.

Conclusion Phase B

This chapter will be closed by concluding with how the DMMAM contributes to the data mesh implementation process. Research from De Bruin et al. (2005), García-Mireles et al. (2012), Lahrmann & Marx (2010), and Wendler (2012) has shown that it is acceptable to map aspects from conventional maturity assessment models to data mesh as these are not restricted to any application domain. At last, it will be concluded which elements will be incorporated into the DMMAM-design.

The provided definition by DAMA International (2017) for maturity assessment referred to the approach to process improvement based on a model that describes how characteristics evolve over maturity levels. Mettler et al. (2010), Schumacher et al. (2016), and Tarhan et al. (2016) mentioned that the maturity assessment model evaluates an organisation's current capabilities and the desirable maturity states to open opportunities for improvement. Al-Sai et al. (2022), DAMA International, De Bruin et al., Korsten et al. (2022), and Lasrado et al. (2015) provided the essential elements of any maturity assessment model, which are the *dimensions*, *characteristics*, and *maturity levels*. In addition, García-Mireles et al. and Wendler explained that each characteristic is defined by specific *criteria and requirements* to distinguish the different maturity levels. According to DAMA International, the criteria and requirements often refer to an activity, tool, standard, and people or resource. Altogether, Becker et al. (2009), De Bruin et al., and Mettler et al. stated that the definitions for the characteristics, aligned over the maturity levels, provide the criteria and requirements as the transformation path towards the complete implementation. Therefore, the DMMAM will provide the data mesh process of implementation.

Three main goals of conducting a maturity assessment were identified by DAMA International. In the context of implementing data mesh, the maturity assessment model will define the path towards complete data mesh implementation. By conducting an assessment, the data mesh activities for the organisation will be discovered and evaluated. Secondly, the stakeholders who are responsible for the data mesh implementation will be educated about the practices, concepts, and principles. In addition, data mesh roles and responsibilities will be identified. Lastly, the maturity assessment will help establish the roadmap in support of implementing data mesh as a strategic goal. Next to the goals, various main drivers for conducting a maturity assessment were identified by Al-Sai et al., Belghith et al. (2021), DAMA International, Kerzner (2001), McCormack et al. (2009), and Lemsa (2021). Reflecting these drivers on data mesh, several findings will be presented. First of all, the maturity assessment could provide regulatory oversight. It functions as a compliance check to assess whether regulatory standards are met. Secondly, the DMMAM could also be used in the form of a readiness assessment. Recognising the need to improve the current practices starts by assessing the current maturity. Thirdly, by asking for both current maturity states and target maturity states in the assessment, the outcomes provide relevant input for planning. Fourthly, the DMMAM would be able to track organisational performance. In other words, the data mesh implementation progress could be assessed and compared to pre-determined goals. The organisation will therefore see whether the adoption is successful. Fifthly, the maturity assessment model could be used as a classification tool, which means that risk, quality, costs, and ROI-values towards the target levels could be assessed. At last, the DMMAM will help as a benchmark tool to see the current progress across organisations within the same industry. Comparing an organisation with its competitors will help establish a competitive advantage. In conclusion, these goals and drivers for using the DMMAM would foster and help the implementation of data mesh.

The benchmark showed that maturity assessment models occur in various forms. By evaluating existing data maturity assessment models, the structure, limitations, and best practices were examined. The benchmark helped define the baseline reference, standards, and provides input for the design and development phase of this research. The models evaluated by Al-Sai et al., Belghith et al., and Król & Zdonek (2020) showed that most models have approximately five maturity levels and five main dimensions. A key conclusion is, based on the statement from Al-Sai et al., that a digital questionnaire, consisting of formulated control assessment questions for the identified elements, supports the application. Al-Sai et al. and Król & Zdonek stated that the form of a self-assessment is desirable such that the organisation could perform the assessment independently. To overcome the limitations as mentioned by Al-Sai et al., several lessons could be learnt, such as providing documentation about the model, explaining the development procedure, identifying dimensions and characteristics, describing the assessment method, presenting a visualisation report with the outcomes, and validating the model by experts and organisations. Belghith et al. compared the strengths and weaknesses of the models. Key strengths referred to offering guidance and details on model features, providing the current state and recommendations for maturity level evolution, providing risks and resource allocation, and providing the opportunity to compare results with other organisations. A weakness is that some models require extensive resources and knowledge. Król & Zdonek noticed that comprehensive descriptions of an organisation's maturity levels were present for all models. However, detailed descriptions of criteria for placing an organisation at a particular development level or about the assessment process itself were often lacking.

At last, design choices will be presented regarding what will be incorporated into the DMMAM-design. It has been concluded that maturity assessment models, regardless of their focus, consist of dimensions, characteristics, maturity levels, and criteria and requirements. Therefore, this will also serve as the starting point for the design of the DMMAM. In addition, approximately five dimensions and five discrete maturity levels will be considered. Furthermore, a best practice referred to providing the assessment as a digital questionnaire in the form of a self-assessment, such that the organisation could perform the assessment independently. Therefore, formulated questions will be added to the characteristics. Moreover, during the design process, special attention will be given to providing documentation about the model, explaining the development procedure, describing the assessment method, offering guidance and details on model features, presenting the results visually alongside the assessment numerical outcomes, and providing detailed descriptions about criteria and requirements across the maturity levels.

Phase C

Design and Development

The design and development phase aims to design the DMMAM. The main sub-question, sub-objective, and sub-deliverable for Phase C are provided below.

Sub-Question 2

What model could be designed to assess the maturity of a data mesh implementation within an organisation?

Sub-Objective 2

Designing the data mesh maturity assessment model.

Sub-Deliverable 2

Developed data mesh maturity assessment model.

In order to provide the answer to sub-question 2, three sub-sub-questions are defined which will be covered in respectively Chapter 4, 5, and 6. Chapter 4 defines a systematic approach to designing the DMMAM to provide the answer to sub-sub-question 2.1. Chapter 5 selects a preferred DMMAM-design. In addition, all model elements will be provided to provide the answer to sub-sub-question 2.2. At last, Chapter 6 defines which outcomes could be obtained from the designed DMMAM to provide the answer to sub-sub-question 2.3.

Sub-Sub-Question 2.1

How to design the data mesh maturity assessment model?

Sub-Sub-Question 2.2

What will be the design of the data mesh maturity assessment model?

Sub-Sub-Question 2.3

What outcomes could be provided by using the data mesh maturity assessment model?

At the end of Phase C, the design of the DMMAM is developed and is ready for demonstration in Phase D.

4 Design Approach

Chapter 4 explains in two sections how the DMMAM will be designed. Section 4.1 introduces a systematic engineering design approach. Section 4.2 will present the phases of the approach. Below are the sub-sub-question, sub-sub-objective, and sub-sub-deliverable for this chapter provided.

Sub-Sub-Question 2.1

How to design the data mesh maturity assessment model?

Sub-Sub-Objective 2.1

Defining a systematic approach to design the data mesh maturity assessment model.

Sub-Sub-Deliverable 2.1

Systematic approach for the design of the data mesh maturity assessment model.

4.1. Systematic Engineering Design Approach

To design the DMMAM, the systematic engineering design process introduced by Dym et al. (2013) will be used. Dym et al. (2013, p. 7) define the design approach as an "*... intelligent process in which engineers generate, evaluate, and specify solutions for devices, systems, or processes whose form(s) and function(s) achieve clients' objectives and users' needs while satisfying a specified set of constraints.*" Using the approach by Dym et al. offers a structured design process. This means that all important design considerations will be addressed in a logical manner. Additionally, this approach puts emphasis on user and client needs. Since the DMMAM will be designed for organisations as potential users and involves Accenture as a client, necessitates considering the user and client's needs and requirements. The relationship between myself as a systems engineering designer, Accenture as the client, and an organisation as the user is incorporated into the systematic engineering design process, making it suitable in the context of this design science research. By following this structure, Dym et al. argue that despite the diverging interests of the designer, client, and user, an applicable and useful model could be designed. At last, Dym et al. also highlighted ethical aspects of engineering design. Designs could also have an impact on people beyond the designer, client, and users. It is important to take into account the ethical implications and ensure that ethical standards are upheld throughout the development process.

Having explained why this approach is expected to be useful, it is worth noting that this approach has already been extensively utilised in other scientific work, which reflects its academic reputation. The systematic engineering design approach is presented in the book *Engineering Design: A Project-Based Introduction*. The first author, Clive L. Dym, was a professor emeritus of Engineering Design and also the director of the Center for Design Education at Harvey Mudd College (Google Scholar, 2023b). His publications have been cited over 12,500 times, reflecting he is well-known in the field of engineering design science. Dym authored and co-authored hundreds of peer-reviewed journal articles, conference proceedings, and technical reports, in addition to publishing thirteen books, including the one in which the systematic engineering design approach is presented. This book has been published by John Wiley & Sons, which is considered an authoritative publishing organisation in fields such as engineering and technology (John Wiley & Sons, 2013). Searching for *Engineering Design: A Project-Based Introduction* on Google Scholar reveals that the 4th edition of the book has been cited over 1,750 times, which is considered relatively frequently. Moreover, the existence of various editions of the book reflects its continued relevance to evolving engineering design practices. Furthermore, his work is also adopted as a course textbook in the master's programme CoSEM. This adoption reflects the credibility of the source and its contribution to the goals of the master's programme, as explained in Section 1.1.2. In addition, this systematic engineering design approach has also been adopted in a book from Alan Hevner, titled *Twelve Theses on Design Science Research in Information Systems*, which has been cited over 18,500 times (Google Scholar, 2023a). Hevner is a distinguished university professor at the University of South Florida. The fact that Dym's work has been adopted by this professor demonstrates the reputableness of the design approach.

On the other hand, when searching *"Maturity Model" OR "Maturity Assessment"* in Google Scholar among the publications that cited Dym et al., only 11 out of the approximately 1,750 results were found. Analysing these publications revealed that only a few addressed topics related to maturity assessment models. However, no publications address maturity assessment models in specific. Despite this approach being expected to be useful for developing the DMMAM, it has not been applied in a similar way before. At last, Leonard et al. (2023) criticise rationalistic approaches, such as presented by Dym et al., in engineering design. Jonassen (2012) states that rationalistic approaches allow engineers to conduct several analyses and compare options to select a single optimal choice. However, Leonard et al. argue that engineering design problems are complex, lacking structure, and are inseparable from socio-cultural contexts, thus requiring diverse decision-making approaches beyond rationalistic methods alone (Jonassen, 2000). Nevertheless, Schön (1983) argues that rationalistic approaches could still be convenient as long as the engineering design problems are also well-structured.

4.2. Systematic Engineering Design Phases

The systematic engineering design process, as defined by Dym et al., will be explained in five phases.

I. Objectives Analysis - The objectives analysis defines the objectives for the design. Dym et al. (2013, p. 7) define a design objective as "*a feature or behavior that we wish the design to have or exhibit.*" The list of objectives could also be presented graphically, by means of a hierarchical objectives tree. An objectives tree is described by Dym et al. as a graphical depiction of the objectives for the device or system.

II. Constraints Analysis – Conducting the constraints analysis results in a list of limits and boundaries the design must meet. Dym et al. (2013, p. 7) define a constraint as "*a limit or restriction on the features or behaviors of the design*" and states that "*... a proposed design is unacceptable if these limits are violated.*"

III. Functional Analysis – The functional analysis shows the functional specifications of the design. Dym et al. (2013, p. 8) define a function as "*things a designed device or system is supposed to do.*" In addition, means will be established for the functions. A mean is defined by Dym et al. (2013, p. 8) as "*a way or a method to make a function happen.*"

IV. Generating Design Space – A morphological chart will be created to present the design space. This chart provides an overview of design choices for the model, based on the established objectives, constraints, functions, and means. According to Dym et al. (2013, p. 25), a morphological chart identifies "*... the ways or means that can be used to make function(s) happen.*" Moreover, Dym et al. (2013, p. 25) state that the morphological chart provides "*... a framework of the design space, an imaginary "space" that we can use to generate potential design alternatives for a design problem.*"

V. Selecting Preferred Design – From the design alternatives, as presented in the morphological chart, a preferred design will be selected. The preferred design is the composition of means for each of the model functions.

By following these five steps, an understanding will be obtained of what is needed from the final model and it shows what model designs will be considered. Chapter 5 continues by conducting these analyses.

5

Model Design

Chapter 5 presents and motivates the design of the DMMAM. The preferred model design will become the outcome of the findings from Chapter 3 and the analyses from the systematic engineering design process as introduced in Chapter 4. Therefore, findings from Chapter 3 will be combined with the insights from the objectives analysis in Section 5.1, constraints analysis in Section 5.2, and functional analysis in Section 5.3. The remaining design choices will be presented in an imaginary design space in Section 5.4. The preferred design will be selected and motivated in Section 5.5. At last, the complete model will be presented in Section 5.6. The sub-sub-question, sub-sub-objective, and sub-sub-deliverable for this chapter are provided below.

Sub-Sub-Question 2.2

What will be the design of the data mesh maturity assessment model?

Sub-Sub-Objective 2.2

Selecting a preferred data mesh maturity assessment model design.

Sub-Sub-Deliverable 2.2

Preferred data mesh maturity assessment model design.

Before proceeding to the design analyses, the information-gathering process will be explained. Dym et al. (2013) state that, in addition to conducting literature research and benchmarking existing maturity assessment models as designer, it is important to involve user and client perspectives in order to inform the design process and acquire the necessary knowledge. It will be explained how these three perspectives will be included.

First of all, *users* refer to the organisations that wish to have their data mesh maturity assessed. In Phase D, the aim is to demonstrate the designed DMMAM in cases. To achieve this, three organisations are identified and approached for participation. Accenture's professional network facilitated the personal outreach to engage these organisations, which are market-leading businesses in financial services and the high-tech manufacturing industry. These organisations were selected as they have started exploring and implementing data mesh. In preparation for the cases, representatives from the organisations were engaged in the design process to gather design knowledge from a user perspective. These interactions specifically focused on discussing the needs and requirements from the user's point of view. These insights were gathered through informal meetings conducted throughout the development period of the DMMAM. To provide the background of the users involved, Table 5.1 presents information about the organisations in terms of the industry in which the organisation operate, the number of employees, the role of the representative who was involved, and the representative's years of working experience. ID-labels are assigned to enable the referencing to the organisations and representatives.

Table 5.1: Organisations and Representatives

	Organisation			***Representative***	
Nr	Industry	Number of Employees	ID	Role	Working Experience
I	Financial Services	20,000	P	Data Officer	10 years
II	Financial Services	40,000	Q	Managing Business Architect	20 years
III	High-Tech Manufacturing	40,000	R	Chief Data Product Owner	10 years

Secondly, Accenture as *client* was involved in the design process by conducting interviews with 15 experts who are employed at Accenture globally. The interviews, which lasted for 45 minutes, were conducted in an informal and semi-structured manner, allowing for a balanced approach of inquiry and conversation. The selection of these experts was based on their knowledge and demonstrated experience with data management, specifically data mesh, as well as their familiarity with maturity assessment models. Table 5.2 presents an overview of the 15 experts in terms of their role within Accenture, the operating group they belong to, and their years of working experience. ID-labels are included to refer to interview participants in the following sections.

Table 5.2: Interview Participants

ID	*Role*	*Operating Group*	*Working Experience*
A	Principal Director	Strategy & Consulting	15 years
B	Senior Manager	Data & AI	10 years
C	Consultant	Data & AI Value Strategy	5 years
D	Analyst	Technology, Strategy, & Advisory	2 years
E	Consultant	Strategy & Consulting	5 years
F	Consultant	Digital Strategy	5 years
G	Managing Director	Data & AI	20 years
H	Manager	Strategy & Consulting	8 years
I	Associate Director	AI Engineering	12 years
J	Associate Director	Technology Data Management	12 years
K	Analyst	Transformation Excellence	2 years
L	Managing Director	Data & Machine Learning	20 years
M	Principal Director	Technology Innovation	15 years
N	Senior Manager	Technology Strategy	10 years
O	Specialist	Technology Data Management	8 years

Appendix B provides more background about the interviews, in terms of the protocol in Section B.1, the participants in Section B.2, the questionnaire in Section B.3, and the responses in Section B.4.

The primary objective of conducting interviews was to discuss the design choices of the DMMAM. In more detail, the following motivations were taken into consideration:

- Exploring the practical experience of data mesh implementations by others.
- Exploring the practical experience of performing maturity assessments by others.
- Exploring the objectives, constraints, and functions of the DMMAM.
- Exploring what else needs to be taken into account while designing the DMMAM.

Thirdly, Dym et al. state that the *designer* (ID: S), incorporates the perspectives of user and client. However, this does not imply that all design decisions are solely based on their needs and requirements. The designer may also have motivations to make additional design choices or approach it differently. Throughout the various analyses, design decisions are typically made by the designer. However, in situations where design choices were motivated by the organisational representatives or interview respondents, the ID-labels will be used accordingly.

5.1. Objectives Analysis

This section presents the objectives for the DMMAM-design. By the designer, the top-level objective has been formulated as aiming for a *successful* DMMAM, which means that the goals of conducting a maturity assessment, as described in Chapter 3, need to be accomplished. As the objectives capture what the design should *be*, the question *"What does a successful DMMAM mean?"* has been discussed in the interviews. The established list of DMMAM-objectives, its interpretation, and the reference(s) accordingly are presented in Table 5.3.

Table 5.3: Objectives Analysis

Nr.	***Objective** (Definition)*	*Reference*
1	**Successful** Accomplishing its goals.	S
2	**Valuable** Important, useful, or beneficial model for the user and client.	B
3	**Well-articulated** Able to express meanings easily and clearly and show their quality.	B
4	**Actionable** Able to be used as a reason for doing something. Outcomes need to be translated into actionable next steps.	B
5	**Tailored** To adjust or expand something to the specific needs of the user and client.	A, B, C
6	**Complete** Data mesh needs to be approached from all the different perspectives.	A, B, C
7	**Feasible** Assessment needs to be able to be performed and it needs to achieve its desired outcomes.	B
8	**Understandable** Able to be understood, so that the user and client know what something means.	D
9	**Client-friendly** Designed from the user and client's point of view. It should meet the needs of the user and client.	D
10	**Reliable** Outcome should be trusted. Important that the client could explain to the user why the final score is reliable.	C
11	**Explainable** Model and assessment outcome should be understood by the user and client.	A, C
12	**Measurable** Aspects need to be measurable to have a correct assessment.	B, C, D
13	**Pragmatic** Solving problems in a sensible way that suits the conditions that really exist now.	B

Nr.	***Objective*** *(Definition)*	*Reference*
14	**Consistent** Always behaving in a similar way. The outcome should be consistent regardless of who the user or client is.	B, C, D
15	**Self-describing** Serving to describe oneself. The user should be able to answer the questions without any help from the client.	A, C
16	**Accurate** Correct, exact, and without any mistakes. Accurate means that it is correct in all the details.	C
17	**Unambiguous** Expressed in a way that makes it completely clear what something means.	A
18	**Recognisable** Concepts need to be familiar to the user and client.	A
19	**Orthogonal** Independent, no overlap in dimensions, characteristics, and maturity levels.	A
20	**Well-defined** Clearly expressed, explained, and described dimensions, characteristics, and maturity levels.	A, B, I
21	**Modular** Consisting of separate parts that, when combined, form a complete whole.	A
22	**Supportive** It should actively give help to the user and client who needs it.	S
23	**Durable** The model should be able to continue to exist for a long time, by being maintainable and sustainable.	S
24	**Convenient** Suitable and comfortable for user and client purposes and needs, and causing the least difficulty.	S
25	**Achievable** An assessment needs to be possible in time and resources.	S
26	**Marketable** Able or fit to be sold or marketed.	S

Table 5.3 shows that a total of 26 objectives have been established for consideration in the development of the DMMAM. Initially, 24 objectives were obtained from the interviews, as presented in Table B.6. However, the objectives *SMART*, *Unbiased*, *Comfortable*, and *Comprehensive* were excluded due to large overlap in interpretation with objectives such as *Measurable*, *Well-articulated*, *Achievable*, *Tailored*, and *Feasible*; *Consistent*; *Convenient*; and *Complete* respectively. Furthermore, the objective *Successful* was added beforehand to state the top-level objective. Additionally, after the interviews, the designer added the objectives *Supportive*, *Durable*, *Convenient*, *Achievable*, and *Marketable*, as these aspects were also deemed important but were not identified during the interviews.

Furthermore, the designer created an objectives tree, as a graphical and hierarchical representation of objectives, based on the list of 26 objectives. Starting with *Successful* as the top-level objective, the tree is further broken down into sub-objectives at different levels to include progressively more detail. Additionally, it clusters objectives that are related to each other. The designer developed this structure iteratively by first examining the various objectives with their definitions, grouping similar and dependent objectives, and creating a quick-and-dirty objectives tree. After making adjustments, the objectives tree for the DMMAM, as presented in Figure 5.1, was finalised. The objectives tree highlights the main trade-off required to achieve success. This trade-off involves maximising the value of the assessment model while maintaining feasibility. To illustrate this trade-off, involving numerous participants in the assessment would increase the likelihood of having profoundly evaluated all data mesh characteristics and thereby obtaining a more accurate maturity score (Loshin, 2008). However, this approach may be infeasible due to limitations in resources such as time, capacity, and budget. Another example relates to the balance between completeness and convenience. If the maturity assessment involves a hundred questions to ensure completeness, it could have a negative impact on the user experience, making the process inconvenient and frustrating (Krosnick, 2018).

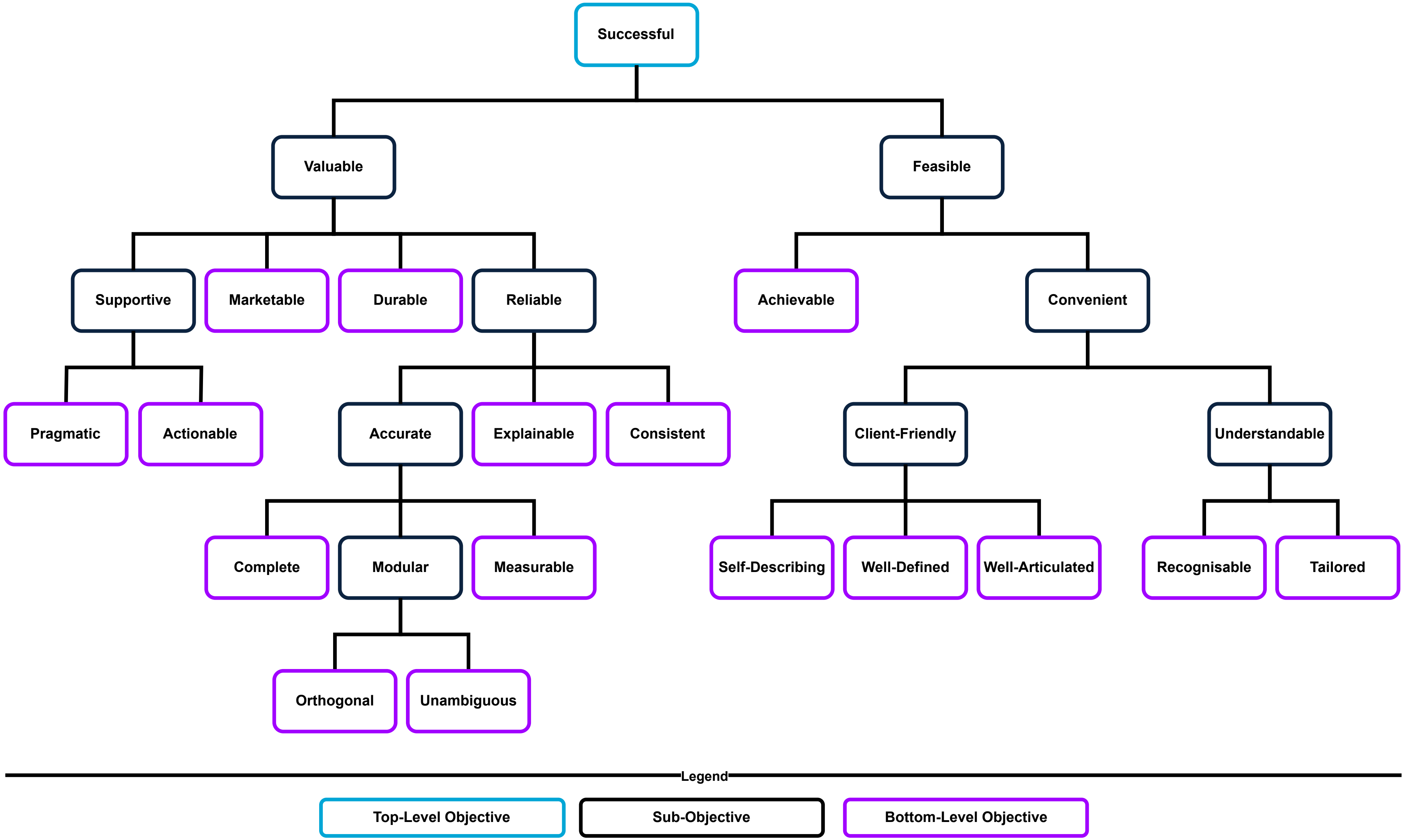


Figure 5.1: Objectives Tree

5.2. Constraints Analysis

This section presents the constraints as the limits and boundaries the design must meet. In the interviews, the participants were asked about the constraints the DMMAM-design must meet. The constraints are subsequently classified by the designer as either model constraints or assessment constraints in Table 5.4, based on the answers presented in Table B.7 and B.17. While model constraints refer to the requirements of the DMMAM, assessment constraints pertain to how the DMMAM needs to be demonstrated in Phase D. It is important to note that although these two aspects are intertwined, for the sake of clarity in the design process, these will be perceived distinctly. The distinction between the assessment model and assessment activities was also emphasised in Section 3.1, with the model elements described in Section 3.1.2, and the activities discussed in Section 3.1.3.

Table 5.4: Constraints Analysis

Nr.	***Model Constraint** (Description)*	*Reference*
1	**Data mesh characteristics must be mutually exclusive and collectively exhaustive:** - All data mesh elements must be covered. - Dimensions and characteristics must be orthogonal, such that there is no overlap.	A, C
2	**Number of maturity levels:** - The model must have at least four different maturity levels. - The model must have no more than five maturity levels.	A
3	**Number of characteristics must balance completeness and research capacity:** - The questionnaire must be able to be completed within 60 minutes. - The questionnaire must not exceed 60 questions. - Each characteristic must have at least one question.	A, D, E, H, N

Nr.	***Assessment Constraint** (Description)*	*Reference*
4	**Number of user participants:** - At least three user participants must be involved. - No more than six user participants must be involved.	A
5	**Represented user participants must be balanced as group:** - Expertise with respect to the different data mesh dimensions. - Number of years experience. - Technical and business stakeholders. - Covering the data mesh supply chain from data producers to data consumers.	B, C, D, E, F, G, N
6	**Duration user participant discussion session:** - Discussion session must be completed in 60 minutes.	A

The constraints presented in Table 5.4 will be explained. Firstly, Participants A and C mentioned that the characteristics need to be mutually exclusive and collectively exhaustive (MECE), meaning that all elements need to be covered without having any overlap. De Bruin et al. (2005) also highlighted the need for having MECE dimensional components and sub-components. Secondly, the number of maturity levels must be at least four states Participant A. To illustrate, having only three levels means that when it is neither classified as *non-initiated* nor *perfect*, it will automatically be at the intermediate level, which makes no sense. On the other hand, having too many levels will result in minimal differences across the levels, which makes it hard for the user to assess where the organisation is positioned. Thirdly, due to limited resources from both client and user perspectives, the questionnaire needs to be completed within 60 minutes (Participants A; D; E; H; N). Expecting that answering one question would on average take one minute, no more than 60 questions are allowed. Fourthly, at least three user participants must be included state Participant A. Due to limited research capacity, it was mentioned that no more than six user participants would be recommended.

This number balances the validity of the obtained final maturity score and the feasibility of the research, which trade-off has become visible in the objectives tree. Fifthly, the group of user participants must be balanced (Participants B; C; D; E; F; G; N). This is needed to guarantee sufficient knowledge is available to answer all the questions which cover the different perspectives of data mesh. The interview participants mentioned that having a balanced group also reduces the risk of having only like-minded people involved, which could result in a one-sided perspective (Holland et al., 1986). Lastly, the designer considered including an assessment activity that involves organising individual discussion sessions with the user participants after they have completed the self-assessment. Participant A mentioned that this session should be completed within 60 minutes, taking into account everyone's limited time and capacity.

5.3. Functional Analysis

This section presents in Table 5.5 the functions and corresponding means to show what the DMMAM is supposed to *do* and how this could be realised. These findings were obtained by asking the respondents about the functions and means of the DMMAM. Moreover, the user perspective was also included by asking what for the organisation is perceived as an important DMMAM-function. Based on the answers from the interview participants, as presented in Table B.8, B.9, B.14, B.15, and B.16, and the user's opinion, four main functions are established by the designer.

Table 5.5: Functional Analysis

Nr.	***Function** (Means)*	*Reference*
1	**Providing overall data mesh maturity score.** 1. Number of maturity levels. 1.1. *Four levels* 1.2. *Five levels* 2. Classification of levels. 2.1. *Numerical levels* 2.2. *Labelled levels* 3. Formula maturity score. 3.1. *Equal weights* 3.2. *Different weights (critical and non-critical elements)*	A, B, C, D, E, F, G, H, I, J, K, L, M, N, O P, Q, R
2	**Providing maturity scores for the different data mesh dimensions.** 1. Number of maturity levels. 1.1. *Four levels* 1.2. *Five levels* 2. Classification of levels. 2.1. *Numerical levels* 2.2. *Labelled levels* 3. Formula maturity score. 3.1. *Equal weights* 3.2. *Different weights (critical and non-critical elements)*	A, B, C, G, J, K, M, N, P
3	**Providing maturity scores from People, Process, Technology perspective.** 1. Number of maturity levels. 1.1. *Four levels* 1.2. *Five levels* 2. Classification of levels. 2.1. *Numerical levels* 2.2. *Labelled levels* 3. Formula maturity score. 3.1. *Equal weights* 3.2. *Different weights (critical and non-critical elements)*	A, B, L, P
4	**Providing guidance for achieving higher levels of maturity.** 1. *Maturity gap* 2. *Prioritising and initiating* 3. *Allocating resources* 4. *Benchmarking* 5. *Achievement benefits* 6. *Urgency to shift*	A, B, D, E, F, H, M, O

Most importantly, according to all interview participants and organisational representatives, the model needs to present an overall data mesh maturity score. Additionally, it is also desirable that maturity scores will be provided for the individual data mesh dimensions (Participants A; B; C; G; J; K; M; N; P). To explain, Participant C highlighted that these sub-scores provide a more insightful outcome. To illustrate, it would be more helpful for the organisation to see where improvements need to be made after the assessment. Next to dimensional scores, scores for the *People*, *Process*, *Technology* (PPT) perspectives, explained by Participant D as the highly adopted *golden triangle* within consultancy, is also recommended to adopt in the DMMAM (Participants A; B; L; P). Moreover, Khodabandeh & Palazzi (1994) and Prodan et al. (2015) state that it is widely recognised that PPT are the three fundamental components that drive improvements in the field of information management. Furthermore, Chen & Popovich (2003) highlight the importance of considering the PPT-aspects when implementing new technological approaches. They emphasise the need for changes in organisational culture to facilitate successful implementations, which encompasses the aspects of PPT. The designer states that the overall score, dimensional scores, and scores for the perspectives could be provided as long as there are different maturity levels, a classification for the levels to address progression, and a formula which determines how the final score will be calculated. According to the conclusion from Chapter 3 that existing maturity assessment models have approximately five levels and the statement by Participant A that the DMMAM must have either 4 or 5 levels, these two options will be considered. In addition, based on the *CMM*, *ISO/IEC 15504*, or *CMMI-DEV* standards, including labels should also be considered to classify the levels and reflect the progression (CMMI Institute, 2012; ISO, 2012; Paulk et al., 1993). According to the formula, Participants C and H explained that dimensions and characteristics have equal weights unless there is a specific reason not to do so. Participant D explained that different weights could be considered to reflect critical and non-critical characteristics. At last, the model design also needs to incorporate the function to provide guidance for achieving higher levels of maturity (Participants A; B; D; E; F; H; M; O). From the interviews, six means were established to provide this guidance, which are the following. Firstly, Participants D, E, F, and G argue that the maturity gap needs to be included as this gap refers to objectively comparing the current maturity score with the target maturity score. This difference would present the starting and endpoint of a potential roadmap for a given scope. Secondly, subjectively prioritising and initiating for which characteristics the maturity needs to be increased opens the process of change (Participants A; B). Participant B motivates that it is important to prioritise and initiate the characteristics after conducting the assessment, as it prevents the outcome from being reduced to just a number. Thirdly, allocating resources means that it is stated what resources are needed to accomplish maturity progress (Participants E; M; O). The participants explained that resources refer to budgets, training, change management programmes, up-scaling capacities, or partnering with organisations. Fourthly, Participant A mentioned the importance of comparing the maturity scores with industry competitors to enable benchmarking. At last, providing insights about what is beneficial and why it is urgent to achieve higher maturity levels would create awareness of why it is needed to change (Participant A). All these functions, specifications, and means will be included in the design space which will be presented in the following section.

5.4. Generating Design Space

The functions, specifications, and means from Section 5.3 are presented in the morphological chart in Figure 5.6. Dym et al. explain that the morphological chart is an imaginary space used to generate the final design. The means as design considerations are provided per specification, from which the preferred means will be selected.

Table 5.6: Morphological Chart

Nr.	*Function*	*Specification*	*Means*	
1	**Overall score**	Levels	*4*	*5*
		Classification	*Numerical*	*Labelled*
		Formula	*Equal weights*	*Unequal weights*
2	**Dimensional score**	Levels	*4*	*5*
		Classification	*Numerical*	*Labelled*
		Formula	*Equal weights*	*Unequal weights*
3	**PPT-score**	Levels	*4*	*5*
		Classification	*Numerical*	*Labelled*
		Formula	*Equal weights*	*Unequal weights*
4	**Providing guidance**		*Maturity gap*	*Prioritising & initiating*
			Allocating resources	*Benchmarking*
			Achievement benefits	*Urgency to shift*

5.5. Selecting Preferred Design

This section will select and motivate from the morphological chart the preferred model design. Figure 5.7 presents the design space with the preferred means selected by the designer.

Table 5.7: Preferred Design

Nr.	*Function*	*Specification*	*Means*	
1	**Overall score**	Levels	*4*	*5*
		Classification	*Numerical*	*Labelled*
		Formula	*Equal weights*	*Unequal weights*
2	**Dimensional score**	Levels	*4*	*5*
		Classification	*Numerical*	*Labelled*
		Formula	*Equal weights*	*Unequal weights*
3	**PPT-score**	Levels	*4*	*5*
		Classification	*Numerical*	*Labelled*
		Formula	*Equal weights*	*Unequal weights*
4	**Providing guidance**		*Maturity gap*	*Prioritising & initiating*
			Allocating resources	*Benchmarking*
			Achievement benefits	*Urgency to shift*

The design choices will be motivated. First of all, the overall score, dimensional score, and PPT-score will have the same amount of maturity levels, classification labels, and formulas to have consistency over the assessment model. Secondly, it has been decided to include four different maturity levels instead of five maturity levels. It is expected by the designer that from four maturity levels, the data mesh implementation process could be provided in a way it would be helpful as guidance for organisations. More specifically, the objectives analysis showed that the top-level objective is obtaining a successful DMMAM, meaning that it aims to achieve the goals as has been concluded in Chapter 3.

It is expected that a four-level DMMAM contributes to discovering and evaluating the characteristics of data mesh, educating the users about the concepts, roles and responsibilities, and it would enable the establishment of roadmaps after the assessment. Moreover, by adding more levels, it would become harder to distinguish the different levels for the designer, user, and client. Furthermore, while emphasising the absence of any DMMAM and the current lack of both academic and empirical research about data mesh as stated by Bode et al. (2023), Butte & Butte (2022), Driessen et al. (2023), Goedegebuure et al. (2023) and Machado et al. (2022), including four maturity levels is perceived by the designer as a solid starting point. In addition, Representatives P, Q, and R expressed the view that organisations are currently exploring data mesh or only having established some first initiatives. Therefore, the designer anticipates that introducing more than four levels at this point would likely create challenges rather than offering clarity in terms of the needed guidance this research aims to provide. Thirdly, the maturity levels will be provided by both a number and label aligned with the standards of maturity assessment models. Due to data mesh being a relatively new concept, many organisations have not even initiated the implementation. Therefore, the first level is level 0, which symbolises the baseline level. From level 0, levels 1, 2, and 3 could be consecutively achieved. The maturity levels are defined as Level 0: *Non-Initiated*, Level 1: *Conceptual*, Level 2: *Defined*, and Level 3: *Achieved*. The level labels are inspired by the *Data Management Capability Assessment Model* (ID: C6) as presented in Table A.4 (EDM Council, 2021). The maturity levels from the EDM Council were selected as a starting point due to the perception that their model is authoritative, as it is an industry-leading data management and governance framework. Additionally, their model includes the *non-initiated* level 0, which sets it apart from the majority of other models that typically start at an *initiated* level 1. Considering that data mesh maturity is expected to be still non-initiated for most organisations, having level 0 as a baseline level is assumed appropriate. Fourthly, the maturity scores will by default be calculated based on equal weights. Assigning different weights to the various dimensions and characteristics is expected to be arbitrary, as the importance of these attributes varies among different user and client perspectives. However, the model design will include an optional function in which critical and non-critical weights could be attached to specific elements. Both the weight levels and distribution of weights over dimensions and characteristics could be adjusted manually. This option has been left out for the PPT-score, due to these perspectives are all equally crucial in data mesh assumes the designer. Fifthly, to provide guidance to the improvement process, the *maturity gap* and *benchmarking* options will be included. The maturity gap will be determined by asking the user participant for filling in both current and target maturity states. Benchmarking is enabled by comparing the maturity scores across organisations by the client. *Prioritising and initiating* is not part of the model design. However, this could be easily obtained by evaluating the results with the user participants. In this research, this has been left out of the scope due to limited time and capacity. *Allocating resources* is hard to incorporate since the relationship between achieving higher maturity levels and the needed resources has not been researched before. The same holds for the *achievement benefits* and the *urgency to shift* means. Having decided on the means for the functions, Section 5.6 will explain how these means will be defined.

5.6. Model Elements

Chapter 3 concluded that the DMMAM will consist of *maturity levels*, *dimensions*, *characteristics*, *questions*, and *criteria and requirements*. The functional analysis in Section 5.3 introduced the PPT-perspectives as an addition. As a result, it has been decided that these six model elements form the structure of the DMMAM. This section further explains in detail how the design elements are established and defined towards the final DMMAM-design.

Figure 5.2 provides an overview of the procedure of developing the six elements while taking into consideration the objectives, constraints, and functions. The study conducted by Lasrado et al. (2015) examined the relationship between the procedure order for defining the model elements and whether the model is intended for a novel or well-established domain. The research discussed two approaches proposed by Mettler et al. (2010), namely the *top-down* and *bottom-up* approaches. According to De Bruin et al., the top-down approach is suitable for relatively new domains as there is little evidence of what maturity is among researchers and practitioners. In contrast, in a well-established domain, the focus lies on the measurement of maturity rather than the definition of maturity itself, requiring the use of the bottom-up approach. Since this research aims to develop the DMMAM for an emerging field, the top-down approach will be adopted. Therefore, the maturity levels will be defined first, followed by the establishment of dimensions, characteristics, perspectives, questions, and criteria and requirements, in that order. Having motivated the order of the design procedure, Section 5.6.1, 5.6.2, and 5.6.3 will present and define the model elements. The final DMMAM is the composition of the six model elements and their elaboration.

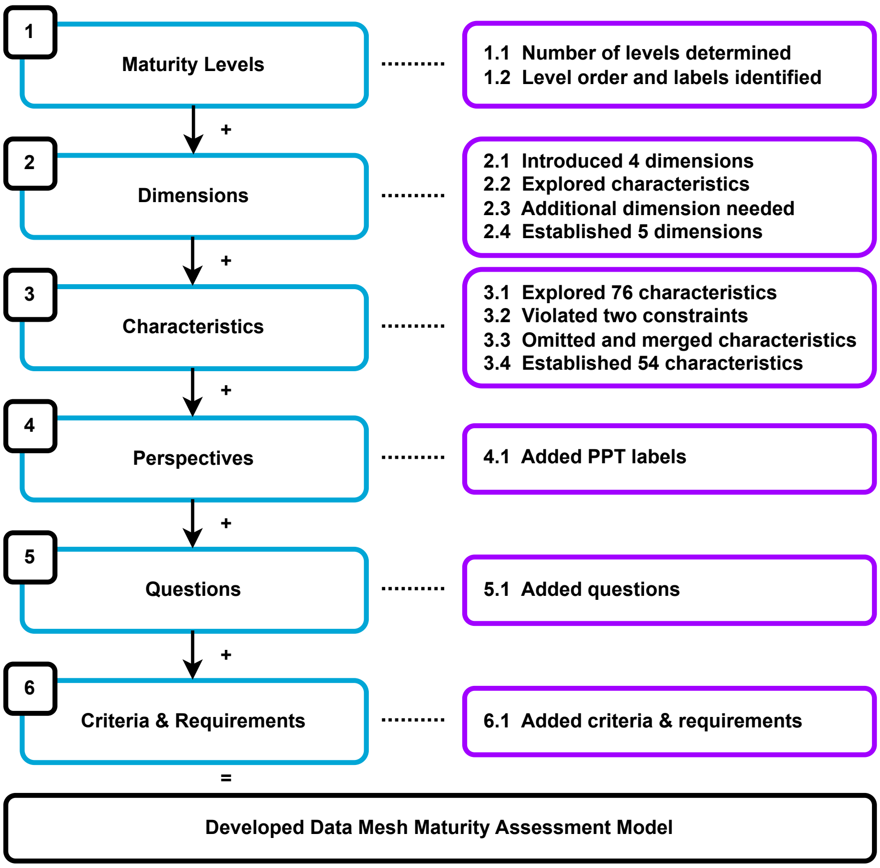


Figure 5.2: Development Procedure Model Elements

5.6.1. Maturity Levels

Section 5.5 decided to include four maturity levels. As the maturity level labels were inspired by the *Data Management Capability Assessment Model* from EDM Council (2021), also their maturity level definitions were used as starting point. According to EDM Council, *Non-Initiated* refers to 'not performed or existing', *Conceptual* refers to 'initial planning stages', *Defined* refers to 'defined and approved', and *Achieved* refers to 'adopted and enforced'. As the *Data Management Capability Assessment Model* has *Enhanced* as the highest maturity level, this definition referring to 'integrated, optimised, and continuously improved' was included in the DMMAM according to the level *Achieved*. The designer has this starting point further elaborated by defining the maturity levels based on three aspects, namely the extent to which the characteristics are *implemented*, the *understanding* of the concept and the *practical experience* with the application of data mesh, as these aspects were also discussed by EDM Council in the explanation of the maturity levels.

The maturity levels, labels, and definitions are provided in Table 5.8. Purple and blue colour scales are added to the maturity levels, which have an identical meanings.

Table 5.8: Maturity Levels

Level	*Label*	*Definition*
0	Non-Initiated	Not having implemented the data mesh characteristics, lacking a comprehensive understanding of the concept, and lacking practical experience with the application of data mesh.
1	Conceptual	Having an initial implementation of the data mesh characteristics, obtained a basic understanding of the concept, and having limited experience with the application of data mesh.
2	Defined	Having established an implementation of data mesh characteristics and obtained a clear understanding of the concept, as well as demonstrable experience with the application of data mesh.
3	Achieved	Having successfully established, optimised, and continuously improved the data mesh characteristics, with a high level of understanding of the concept, and the organisation is highly experienced with the application of data mesh.

5.6.2. Dimensions

Data mesh is coined by Dehghani (2022a) as the convergence of four main dimensions, which are *domain oriented decentralised data ownership and architecture*, *data as a product*, *self-serve data infrastructure as a platform*, and *federated computational governance*. In the interviews is by Participants A, B, C, G, J, K, M, and N and Representative Q proposed to use these four dimensions as starting point. This would make it convenient for the user and client since these dimensions will be directly recognised as the data mesh dimensions from Dehghani. Subsequently, characteristics were explored in the interview sessions and by examining literature, aiming for dividing these characteristics into the four dimensions.

According to Dehghani, it is needed to have a data foundation in place to be able to implement data mesh. In addition, implementing data mesh asks for organisational change. However, it is hard to divide these generic data management, people, and cultural characteristics over the four dimensions. Since data mesh is perceived as an approach to data management, Representative Q mentioned that it may be valuable to add an upper layer of data management characteristics that is crucial for the functioning of the four data mesh principles. Therefore, it has been decided in this research to add a fifth dimension which covers the data foundation and organisational change aspects. Figure 5.3 shows the established dimensions for the model and the relationship between them. Dimension A is considered the baseline with respect to dimensions B, C, D, and E. This means that the data foundation and organisational change characteristics will also be applicable as starting point for the other dimensions. In other words, it provides the foundation to start implementing the four core dimensions.

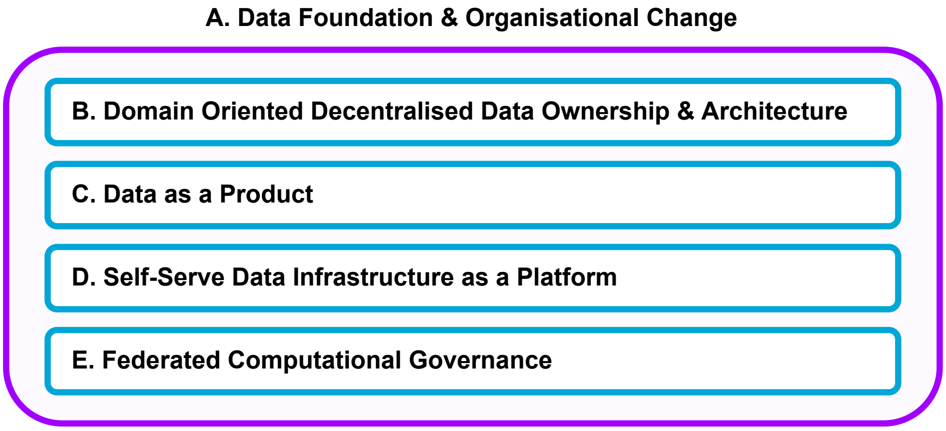


Figure 5.3: Model Dimensions

To have a common understanding of the dimensions, definitions will be given in Table 5.9. The definition for dimension A is established to refer to the needed organisational data and technology fundamentals and to incorporate that Dehghani (2022a, chap. 16, para. 1) states that data mesh introduces changes to "*... people's roles, responsibilities, motivations, and collective interactions in an organization.*" The definitions for dimensions B, C, D, and E align with the descriptions of the dimensions as explained in Section 2.1.3.

Table 5.9: Model Dimensions

ID	*Dimension*	*Definition*
A	Data Foundation & Organisational Change	Organisational data fundamentals and the needed changes introduced to people's roles, responsibilities, motivations, and collective interactions in an organisation.
B	Domain Oriented Decentralised Data Ownership & Architecture	The ownership of analytical data is decentralised to business domains closest to the data. The data-sharing responsibilities should lie with those who are most familiar with the data.
C	Data as a Product	Analytical data provided by the domains is treated as a product. Applying product thinking to how data is modelled and shared to serve business use cases. Ensuring data products meet all usability attributes.
D	Self-Serve Data Infrastructure as a Platform	Data infrastructure which integrates operational and analytical capabilities into a self-serve data platform to empower domain's cross-functional teams, with decentralised technologies, to build and share interoperable data products autonomously to serve domain-agnostic use cases.
E	Federated Computational Governance	Decision-making model that balances autonomy, agility, and local decision-making power of domains, while creating and adhering to defined global rules. The governance execution model relies on automation by computational policies, for every data product, via the platform services, to assure data is secure, compliant, of quality, and usable.

5.6.3. Characteristics

This section will introduce the characteristics of each of the five dimensions. Section 1.1.1 described that Dehghani is acknowledged globally by researchers and practitioners as the founder of data mesh. Moreover, Participant J emphasises that her book *Data Mesh* is currently the most reputable source to learn about data mesh. In other words, her book is perceived as authoritative as it provides a comprehensive in-depth explanation of what data mesh entails in contrast to plenty of grey literature. Moreover, due to the limited availability of academic publications and empirical research on data mesh, which has been a focal point of this research, the book by Dehghani is analysed and used in conjunction with interview responses as the primary input for establishing the characteristics.

By conducting interviews and reading the book, 76 characteristics were found initially. These characteristics are presented in Table C.1 from Appendix C based on the interview participant responses as presented in Table B.10, B.11, B.12, and B.13, and from reading the book from Dehghani. However, this extensive set of characteristics violated two of the constraints from Section 5.2. Given each characteristic has one question, and one question will take one minute, 76 characteristics exceed the time constraint of completing the assessment within 60 minutes. In addition, the established characteristics were also not mutually exclusive, implying that there was overlap assumed. Violating these two constraints required revision.

Characteristics were merged or omitted by the designer when it became evident that certain characteristics were closely related to each other or did not directly relate to data mesh particularly. This revision resulted in a reduction of 22 characteristics. Table 5.10 provides several examples, which will be explained, of how characteristics were merged to reduce the overall number while maintaining the completeness of the DMMAM. First of all, an established A4: *Vision* provides the baseline for developing an A1: *Data-Oriented Strategy* and could also be interpreted together as strategic vision according to Morris (1987), which means these characteristics are closely related. Secondly, A12: *Culture & Mindset*, A15: *Values*, and A20: *Awareness Importance* all focus on enabling a culture and mindset in which people understand the importance of data and take actions lived through defined values according to Dehghani. Thirdly, A16: *DevOps*, A17: *DataOps*, and A18: *MLOps* are all linked to A5: *Agile*, since these methodologies emphasise collaboration, automation, and continuous delivery (Atwal, 2020). Fourthly, A8: *Skills & Capabilities* refer according to Dehghani to data mesh literacy, which means that all people in the organisation are able to perform their data mesh activities and are experts in their own data. This includes that the goal of A13: *Democratisation* is achieved (Bandari, 2020). Fifthly, C6: *Publication* and C5: *Sharing* are closely linked to each other, since Dehghani explains that produced data products serve cross-domain use cases in a data mesh. In other words, the produced data products are developed to be shared. Sixthly, E4: *Security* and E1: *Compliance* address the need for defined security policies and having security governance tools in place to check compliance (Chen & Popovich, 2003). Lastly, E10: *Computational Decision-Making*, E14: *Data Governance Automation*, and C3: *Embedded Governance* all focus on embedding policies as automated and machine-led processes to enable automation (Rahimzadeh et al., 2022). Eventually, the second iteration resulted in a list of 54 characteristics.

Table 5.10: Merging Characteristics

Initial ID	Initial Characteristic	Revised ID	Revised Characteristic
A1 A4	Data-Oriented Strategy Vision	A1	Data-Oriented Strategy & Vision
A12 A15 A20	Culture & Mindset Values Awareness Importance	A2	Culture, Mindset, & Values
A16 A17 A18	DevOps DataOps MLOps	A5	Agile
A8 A13	Skills & Capabilities Democratisation	A11	Skills & Capabilities
C6 C5	Publication Sharing	C4	Production & Sharing
E1 E4	Compliance Security	E1	Security & Compliance
E10 E14 C3	Computational Decision-Making Data Governance Automation Embedded Governance	E6	Computational Policies & Automation

During the process of discovering all the characteristics and developing the final set, three reviewers from Accenture, as outlined in Table 5.11 as Reviewers T, U, and V, were involved individually in 60 minutes discussion sessions.

Table 5.11: Reviewers

ID	*Role*	*Operating Group*	*Working Experience*
T	Manager	Data & Technology	10 years
U	Senior Manager	Data & AI	20 years
V	Managing Director	Data & AI	20 years

In addition, Representatives T, U, and V were also informed during both iterations of developing the characteristics. However, very few to no changes resulted from these engagements, as everyone agreed that this list captures the essence of data mesh. The involvement of these individuals highlights the expertise they brought to the table, prepared to make any necessary adjustments if needed, although ultimately it was not required. Nevertheless, these sessions greatly aided me as a designer, as they compelled me to explain everything clearly. This highlights the importance for designers to have a thorough understanding of the intended meaning behind specific characteristics and how to effectively communicate their definitions. Ultimately, the set of 54 characteristics was not further revised. The final set of characteristics will be presented at the end of this chapter.

People, Process, Technology Perspectives

For the established 54 characteristics, PTT-labels were assigned. Descriptions from Prodan et al. are used for defining the PPT-perspectives in this research, which will be explained. People refer to individuals who possess the necessary expertise, skills, motivation, and collaborative engagement to successfully perform activities. Process implies the series of related tasks that are performed in a particular order, where each task uses specific inputs and adds value to create desired outputs. At last, technology is described as the tools, techniques, hardware, and software utilised for communication and enhancing work efficiency, such as information management systems and their architectures. Table 5.12 shows what is considered by the PPT-perspectives in this research.

Table 5.12: People, Process, Technology Perspectives

Label	*Definition*
People	Human resources including their skills, expertise, and knowledge and the collaboration between them.
Process	Policies, procedures, and workflows to carry out tasks in a systematic and standardised way.
Technology	Tools, software, and hardware used to support organisational operations.

Questions

Questions will be formulated for each of the 54 characteristics. The purpose of these questions is to provide a description of how each characteristic will be measured and enable the self-assessment by the participants.

Criteria and Requirements

Criteria and requirements, representing an activity, tool, standard, and people or resource, will be established following the progression as reflected by the maturity labels. The process of establishing descriptions for all 54 characteristics across four maturity levels involves a total of 216 descriptions that need to be defined. While Appendix D provides a more detailed description of this process, key steps will be highlighted below.

The designer gathered information from the literature and interviews on all 54 characteristics, encompassing relevant details that could contribute to describing each characteristic. In an iterative manner, draft versions of descriptions were created for all characteristics. Everything important for measuring the maturity of a characteristic was documented. It was observed that some characteristics had more uniform descriptions, while others exhibited a wide range of diverse aspects. After drafting all the descriptions, an attempt was made to obtain the essence of each one. The core of each description became the definition of the characteristic. Using this definition as a basis, the next step was to determine how the classification could be made in relation to the four levels of maturity. The initial focus was on Level 0: *Non-Initiated* and Level 3: *Achieved*, as they represent the states of non-implementation and perfection, respectively. Formulating the question required careful consideration, as the descriptions of different maturity levels for a characteristic provided multiple-choice answers. This indicates the iterative nature of the process. When establishing the descriptions for the criteria and requirements, the activities, tools, standards, and people or resources related to the characteristic were considered. Additionally, the definitions of the various maturity levels presented in Table 5.8 were taken into account while creating the categorisation for the different levels. After establishing the descriptions for Level 0: *Non-Initiated* and Level 3: *Achieved*, Level 1: *Conceptual* and Level 2: *Defined* were also filled in. Altogether, this process was conducted for all 54 characteristics, which indicates that this was a part of the research that required a lot of time and effort.

For all dimensions, the characteristics with corresponding questions, perspective(s), and maturity level criteria and requirements will be presented. In addition, the references for the interview participants have been provided using their respective ID's, and the book from Dehghani is referred to as Reference Z. To emphasise the different dimensions, purple and blue colour scales are alternating.

Dimension A: Data Foundation & Organisational Change

A1. Data-Oriented Strategy & Vision

Defined an effective data strategy for how the organisation takes advantage of data & analytics and machine learning as strategic differentiators. Established vision to use data, which defines baseline or "North Star" (desired future end-state) for data strategy.

Question: *Is your organisation adopting a data-oriented strategy, based on a vision to use data?*

Level 0	Level 1	Level 2	Level 3
No effective data-oriented strategy and vision considered. Lack of understanding of how data will drive business growth. No roadmaps defined for the development of data initiatives.	Initial data-oriented strategy and vision created. Ad hoc data & analytics initiatives developed. Often not considered to be aligned with the initially created data strategy. Target states not determined.	Comprehensive data-oriented strategy and vision in place. Data & analytics and machine learning initiatives are developed where organisational data strategy is backing the decisions. Data initiatives serve the overall business interest. Roadmaps were created to achieve target states.	Effective data-oriented strategy and vision adopted. Data is treated as strategic asset, which directly drives business growth. Strategy drives gaining competitive advantage by using data & analytics and machine learning as differentiators. Current state assessments are performed to create roadmaps to achieve target states for all functions.

Perspectives: People, Process, Technology Reference: Z

A2. Culture, Mindset, & Values

Enabled self-serve culture in which people understand the importance of data, are data literate, have a product-sharing mindset, trust the validity of data and accomplish data use cases. The actions and decisions teams and individuals make are lived through defined values.

Question: *Does your organisation has a data culture with defined values where people apply their data product-thinking mindset?*

Level 0	Level 1	Level 2	Level 3
Data as an asset to be collected and protected within the domains. People do not understand the importance of data usage and sharing. Lack of common language, training, and communication. No data literacy. No values cultivated to establish data culture.	Cross-domain data sharing initiatives set up. Only a few people understand how data supports business operations and future initiatives. Only a few people have a common language, attend training, and communicate. Values defined and underpinning data culture.	Real data product-thinking mindset. Data is used and shared to power the operations and drive innovation. People are data literate and attend training. Values defined and cultivated to establish data culture.	Cultural shift from data protection towards data sharing accomplished. Data as a product thinking serves cross-domain business use cases. Culture in which people are data literate and trust the validity of the data. Defined values are cultivated in the organisation. Values live through the actions and decisions teams and individuals make.

Perspective: People Reference: I, L, M, O, Z

A3. Value Realisation

Offered and tracked distinct value by data & analytics and machine learning.

Question: *Is your organisation offering and tracking distinct value to customers and partners using data & analytics and machine learning?*

Level 0	Level 1	Level 2	Level 3
No or limited business value realised by data & analytics and machine learning. Realised value is only examined at the start of an initiative. No or few metrics are established to track the progress.	Some projects identified which obtain business value by data & analytics and machine learning. Throughout the programme, metrics are captured which provide insights into the progress. During the programme life-cycle, improvements are considered manually.	Continuous value realisation by data & analytics and machine learning initiatives. Progress is tracked by monitoring the performance metrics and by including alerting triggers. Tooling provides real-time insights into where enhancements are needed.	Value realisation targets set and being hit both financially and strategically by using data & analytics and machine learning. Value target progress is tracked throughout the process. Dashboarding and platform tooling provide insights into progress. Continuous alerting and enhancement recommendations provided to feed back into decision-making.

Perspectives: Process, Technology Reference: Z

A4. Curiosity & Ability

Established ubiquitous culture of data curiosity and experimentation to build and create the technology needed to embed data sharing and consumption at the core of each business function.

Question: *Is your organisation curious about data usage and experimentation as a foundation for implementing their strategy?*

Level 0	Level 1	Level 2	Level 3
No curiosity for using data and experimentation or only the central IT-team takes responsibility to use and experiment with data.	Only a few organisational domains are using and experimenting with data. These domains have an appreciation for meaningful, trustworthy, and secure data, whereas other domains are not participating. Most domains are not able to experiment with data.	Organisational-wide interest in data usage and experimentation. Almost all organisational domains have meaningful, trustworthy, and secure data. Data is however protected within domains, instead of cross-domain data sharing.	Ubiquitous culture of data curiosity and experimentation. Organisational-wide enthusiasm around data usage and learning. A culture that obsessively runs data experiments, observes the results, analyses the data, makes sense of it, learns from it, adapts, and shares it in a meaningful, trustworthy, and secure way.

Perspectives: People, Technology Reference: Z

A5. Agile

Developed working environment that is iterative in nature and highly collaborative between different teams. Practices such as DevOps, DataOps, and MLOps are integrated.

Question: *Is your organisation experimenting with the data mesh implementation by an iterative approach to incrementally move forward?*

Level 0	Level 1	Level 2	Level 3
Agile methodology and iterative delivery not reflected in ways of working in teams. Domains work in silos and do not collaborate.	Only few domains use agile methodology. Not consistently applied throughout the organisation. Limited infrastructure to support DevOps, DataOps, or MLOps techniques.	Standardised DevOps, DataOps, and MLOps processes. Ways of working aligned with agile principles. Agile tools help process improvement. The methodology is applied to almost all domains in the organisation.	Developed strategy based on iterative delivery and agile methodology. The working environment is highly collaborative between teams. Embedded processes are integrated to track and measure change. Practices such as DevOps, DataOps, and MLOps are integrated as part of the organisation's operating model. Agile tools help continuous process improvement.

Perspectives: People, Process, Technology Reference: Z

A6. Executive Commitment

Secured executive support and top-down engagement of leaders (CDO-, CTO-, CIO-offices).

Question: *Does your organisation has support for data initiatives, such as data mesh, by C-level executives?*

Level 0	Level 1	Level 2	Level 3
No C-level recognition that data is important in your organisation. No budget available to invest in new data capabilities.	Leaders do recognise the business value of data. However, no C-level executive support or only CDO recognises the importance. Financial support is only provided ad hoc.	Data initiatives are continuously supported by leaders and receive funding. Technical C-level executives drive data strategy to the CEO's agenda.	Continuous top-down communication and long-term executive support to enable data initiatives. Actively investing in needed resources by leaders to hit strategic targets. C-level executives are highly involved in projects and their progress. All leaders recognise the importance of data.

Perspectives: People, Process Reference: C, Z

A7. Solid Engineering

Established solid engineering practices and access to modern data tooling.

Question: *Does your organisation has modern engineering practices in place as a foundation to bootstrap data mesh?*

Level 0	Level 1	Level 2	Level 3
No engineering practices implemented. No access to open and modern data tooling.	Limited engineering foundation in place. No or limited continuous and automated delivery of software, DevOps practices, distributed architecture, computational policies, data storage, and processing stacks on the cloud. Hardly able to build data-driven technology to enhance business.	Well-established engineering foundation implemented. Continuous and automated delivery of software, DevOps practices, distributed architecture, computational policies, data storage, and processing stacks on the cloud. However, hardly able to build data-driven technology on their own initiative.	Solid engineering foundation established. Continuous and automated delivery of software, DevOps practices, distributed architecture, computational policies, data storage, and processing stacks on the cloud. Ability and desire to build data-driven technology to enhance business by own initiative.

Perspective: Technology Reference: Z

A8. Change Management

Prepared initiatives to accomplish cultural, organisational, and technological change.

Question: *Is your organisation setting up change management programmes to facilitate the migration from traditional data management approach towards data mesh?*

Level 0	Level 1	Level 2	Level 3
No change management programmes in place to drive organisational transformation.	Change management programmes initiated. Employees' behaviour, attitude, and capabilities will be aligned according to data mesh principles. However, programmes are temporary, not organisational-wide focused, and merely focused on specific data mesh aspects.	Change management programmes organised and continuously provided to facilitate the migration towards a data mesh. Since data mesh will be implemented incrementally, only a few aspects have been covered yet. Clear bottom-up enablement through technology, incentives, and education provided.	Continuous provision of programmes to stimulate and facilitate the complete migration towards a data mesh, consisting of the formation of a federated governance operating model, formation of the cross-functional business, dev, data, ops-teams, and establishment of the data mesh roles. Bottom-up enablement through technology, incentives, and education.

Perspectives: People, Process, Technology Reference: G, O, Z

A9. Value Adding Use Cases

Identified use cases for data products to create inherent value for the data users in service of the business and customers.

Question: *Are there in your organisation business initiatives identified where data products will create direct value?*

Level 0	Level 1	Level 2	Level 3
No use cases identified. Data as a product does not exist.	Data-oriented business initiatives are identified. The identified use cases for data products are shared ad hoc across domains, however with limited usage of governance and the self-serve platform.	Business use cases driven by data & analytics and machine learning serve as the means to identify and deliver data products, enable domains to adopt data, and establish governance and the platform.	Machine learning and analytics-powered business use cases become the vehicles that execute the identification and delivery of data products, adoption of data by domains, and establishment of governance and the platform. In addition, all data products are having unique value, so no duplication.

Perspectives: Process, Technology Reference: Z

A10. Roles

Defined data mesh roles for the people in the producer, consumer, governance, and platform teams who perform the tasks.

Question: *Are in your organisation the producer, consumer, governance, and platform roles assigned?*

Level 0	Level 1	Level 2	Level 3
No roles defined and assigned.	Only few roles initiated and currently assigned.	Most of the roles are identified and assigned to them responsible for the producing, consuming, governance, and platform tasks.	For the people who perform tasks within the data mesh, all roles are clearly defined and assigned. All roles are identified in the data domain-based producer and consumer teams, central data governance team, and central self-service data infrastructure platform team.

Perspective: People Reference: C, I, O, Z

A11. Skills & Capabilities

Needed business, technology, and data skills and capabilities (data mesh literacy) for people actively participating in the cross-functional domain, governance and platform team to perform their activities and for all other people across the organisation regardless their role and level.

Question: *Are people in your organisation have the right skills and capabilities to perform their key data mesh activities?*

Level 0	Level 1	Level 2	Level 3
Lack in skills and capabilities to perform the required activities in the data mesh.	People participating in the data mesh are able to perform their activities correctly. Overall, data mesh literacy is however low.	All participating people understand their responsibilities in the data mesh. Despite almost all data mesh roles are assigned, people need to become more data mesh literate in general. People understand and feel responsible for their own data.	All people are data mesh literate and are able to correctly perform the activities according to their assigned (ownership) roles (business, technology, data, platform, governance). All people in the data mesh are experts in their own data. The goal of data democratisation is achieved.

Perspective: People Reference: Z

A12. Incentivisation

Defined and integrated motivations and rewards that align goals of teams and individuals with the overall data strategy, based on data mesh ways of working.

Question: *Are incentivisation mechanisms defined in your organisation based on domain autonomy, purpose, and progress to rewarding teams or individuals?*

Level 0	Level 1	Level 2	Level 3
No incentivisation mechanism available.	Motivations and rewards provided only ad hoc. No metrics or clearly defined rules are identified.	Frequently provided motivations and rewards. Clearly defined metrics of how performance is measured for teams and individuals. However, current rewards are less incentivising.	Motivations and rewards integrated that align goals of teams and individuals with the overall data strategy, based on data mesh ways of working - autonomous teams delivering results and growing a new generation of data practitioners. Metrics monitored and a wide range of rewards defined: bonuses, incentive plans, ranking, promotions, working conditions, analytical budget etc.

Perspectives: People, Process Reference: G, I, K, Z

A13. Training

Organised educational programmes to create awareness of what data mesh is, why it is relevant to the organisation, getting acquainted with new tools, and learning software engineering.

Question: *Are in your organisation educational programmes in place to enable people to successfully execute the data mesh?*

Level 0	Level 1	Level 2	Level 3
No or limited data mesh education programmes in place. People in the organisation do not have an understanding of what good data is and how it could be utilised.	Trainings are sporadically organised. Data awareness is present across the organisation, but not everyone is actively working with data. Training (data literacy, technical hard skills, appreciation programmes) would be needed to further enhance data mesh implementation.	Organisation has standardised training process. Knowledge exchange tools are available to gain the most relevant information. These programmes really help people to gain knowledge to better perform their activities.	Educational programmes have created awareness, appreciation, empathy, and curiosity about data for everyone. Personal skill development programmes live for continuous learning (e.g., executives get data literacy training and technologists/data specialists/platform engineers get acquainted with new tooling, software engineering, or building/hosting applications.

Perspectives: People, Process, Technology Reference: Z

Dimension B: Domain Oriented Decentralised Data Ownership & Architecture

B1. Definition

Defined what a domain is within the organisation (team, group, tribe, business unit, etc.)

Question: *Is in your organisation defined what is considered by a domain?*

Level 0	Level 1	Level 2	Level 3
No agreed definition of what a data domain is in the organisation.	The few established domains are set up without definition or only with a vague definition. Domains are ad hoc created.	Definition exists for specific domains, but does not logically follow an organisational-wide definition. Only for the current participating domains, a reasoning is established why this domain structure has been chosen.	Clear definition of what a domain is. E.g., domain definition logically follows team, value stream, or business unit structure or business domains are divided based on its resources, responsibilities, outcomes, and knowledge.

Perspective: Process Reference: Z

B2. Structure

Defined structure of how domains are organised and how domain-driven thinking is integrated.

Question: *Is your organisation structured into autonomous domains?*

Level 0	Level 1	Level 2	Level 3
Organisation is not structured into autonomously operating domains.	Small number of autonomous domains identified.	Organisation is almost completely structured into autonomous domains. However, not all domains have been mobilised yet.	Sustained number of autonomous domains defined. Domain-driven thinking is completely integrated across the organisation.

Perspectives: People, Process Reference: G, L, Z

B3. Decentralisation

Distributed responsibility and accountability to domains which are most familiar with the data.

Question: *Are the data responsibilities decentralised to domain-level?*

Level 0	Level 1	Level 2	Level 3
Centralised ownership. Data responsibilities are allocated centrally by the IT or Analytics team.	Most responsibilities are covered by the central IT or Analytics team. However, some business units which for example have close connections to their data sources take responsibility and feel accountable.	Most responsibilities are decentralised into the domains. Ownership and responsibility of the data life-cycle are within the originating business domains. Central IT or Analytics team is still supporting the domains.	All data sharing responsibilities lie at those domains which are most familiar with the data. Decentralisation and distribution of responsibility support continuous change and scalability. Life-cycle of domain-oriented data is managed independently of a central IT or Analytics team.

Perspectives: People, Process Reference: Z

B4. Ownership

Defined and allocated domain owner.

Question: *Are in your organisation domain owners defined and allocated to take responsibility and accountability for the domain data?*

Level 0	Level 1	Level 2	Level 3
Neither domains nor domain owners defined. In case domains are identified, no domain owners are defined.	Some domains have a clear domain owner. There are also domains without clear domain ownership.	For all established domains are domain owners allocated. Domain owners take responsibility and accountability for domain data.	Ownership exists and is well-defined for all domains in the organisation. Domain owner has the resources to act and innovate. Domain owner focuses on the data which is shared by that domain and takes full responsibility and accountability for the domain data.

Perspectives: People, Process Reference: C, M, O, Z

B5. Autonomy

Enabled autonomy for domains by having the operational latitude, team structure, and skills necessary to manage their data independently.

Question: *Are the domains in your organisation operating fully autonomous?*

Level 0	Level 1	Level 2	Level 3
Non-autonomous domains. Domains are not able to self-govern in the data mesh.	Domains are to some extent able to perform data activities on their own. However, still dependent on a central IT or Analytics team with, for example, data product building, using the platform, integrating governance in data products, and assigning data product ownership.	Domains are able to perform their data activities almost independently, and in general, no support is needed from the central IT or Analytics team. Cross-functional teams have the skills to perform the activities autonomously. Support from central teams is only needed sporadically.	All domains are operating fully autonomous. Domains have direct ownership of their data and have the latitude, cross-functional team structure, and skills necessary to perform all activities. Each domain has complete autonomy in serving its operational and analytical data assets.

Perspectives: People, Process Reference: K, Z

B6. Cross-Functional Teams

Established cross-functional teams within the domains (business, technology, data).

Question: *Are the domains internally balanced in terms of teams or individuals with business, technology, and data skills & capabilities?*

Level 0	Level 1	Level 2	Level 3
Domains only have functional teams.	Some domains have clear cross-functional teams. In other domains, not all skills & capabilities are covered to autonomously use and share data by self-serve tooling. Still dependent on, for example, central IT or Analytics team to perform activities.	Almost all domains have cross-functional teams. Business, technology, and data skills & capabilities are covered within the domains. Only in a few cases, external support is needed to support the domain with functional knowledge.	Within all domains, cross-functional business, technology, and data teams are created each responsible for long-term ownership of their data. Alignment of technology teams working in close collaboration with their business counterparts. Cross-functional collaboration enables cross-pollination of different skill sets. No external support is needed.

Perspective: People Reference: Z

B7. Architecture

Distributed architecture matching the domain-oriented organisational autonomy where data products are accessed through standardised protocols.

Question: *Is the organisational architecture matching the distribution of domains?*

Level 0	Level 1	Level 2	Level 3
No distributed architecture. Data architecture is more monolithically organised with a central data warehouse or data lake.	Data architecture is in general still monolithic. However, an initial domain-oriented architecture is established for some domains where they stored and managed their data decentrally.	Most domains are covered in the decentralised data architecture. These domains are having own infrastructure technologies which enable them to decentrally store and manage data.	The architecture completely matches the domain-oriented organisational autonomy, with a corresponding data-product-oriented distributed architecture. The complete architecture migrated from a monolithic towards a completely decentralised architecture.

Perspectives: People, Process, Technology Reference: Z

B8. Producers

Participating domains producing data products.

Question: *How much domains are providing analytical data to other domains?*

Level 0	Level 1	Level 2	Level 3
No domains are producing data products. Data production is an IT-responsibility.	Some domains have started building data products.	Almost all domains are producing data products. Continuous supply of data products to be shared for cross-domain use cases.	All domains are producing data products and do this frequently. There is domain dependency on data products. Positive network effect created by peer-to-peer connectivity of domains exchanging data products as units of value.

Perspectives: People, Process Reference: Z

B9. Consumers

Participating domains consuming data products.

Question: *How much domains are consuming analytical data from other domains?*

Level 0	Level 1	Level 2	Level 3
No domains are consuming data products. Data consumption is an IT-responsibility.	Some domains have started consuming data products.	Almost all domains are consuming data products. Continuous demand for data products to be collected for use cases.	All domains are consuming data products and do this frequently. Domains have dependencies on data coming from one or more other domains.

Perspectives: People, Process Reference: Z

Dimension C: Data as a Product

C1. Definition

Defined what a data product and data as a product are within the organisation.

Question: *Is in your organisation defined what is considered by a data product and data as a product?*

Level 0	Level 1	Level 2	Level 3
No agreed definition on what a data product and data as a product are within the organisation.	Understanding within the organisation what a data product and data as a product are. However, this is only vaguely defined by different domains.	Definition exists for data product and data as a product in the domains, but does not logically follow an organisational-wide standardised definition.	Organisational-wide clear standardised definition of what a data product and data as a product are. E.g., defined as a unit of the logical architecture, a data quantum, controlling and encapsulating all the structural components needed to share data as a product to serve use cases.

Perspective: Process Reference: Z

C2. Ownership

Establishment of the data product ownership role, defined data product owner, and taking ownership of data product responsibilities.

Question: *Are the data products have a data product owner who takes the data product responsibilities?*

Level 0	Level 1	Level 2	Level 3
Data is an IT-owned asset. IT or Analytics team is responsible for the data. No data product owners in the domains allocated.	Only in some cases data product owners are allocated. No clear data product owner allocation governance process in place. Often unclarity or discussion about who the data product owner will be.	Most data products have a data product owner who takes the responsibility to provide understandable and usable data products to other domains. However, not all data products have data product ownership defined. A clear understanding is obtained of who the data product owner will be after producing a data product.	All data products have an assigned data product owner. Data product owners take the long-term ownership of responsibilities to create, model, maintain, evolve, and share data as a product to meet the needs of data product consumers. Domain's data product owner must ensure data is delivered to certain KPI's, such as quality, adoptability, usability and security.

Perspectives: People, Process Reference: I, N, O, Z

C3. Discovery Tool

Established global data product discovery tool or inventory (data catalogue/marketplace) where data products are automatically accessible. The tool enables a seamless way to search, explore, request, and share relevant data products.

Question: *Is there in your organisation a data discovery tool available which enables searching, exploring, requesting, and sharing data products?*

Level 0	Level 1	Level 2	Level 3
No data discovery tool in place. It is not possible to search, explore, request, or share data products.	Data discovery tool in place. However, difficult process to find and share data products.	Data products are accessible through the data discovery tool. The tool enables the business to easily search, explore, request, and share data products.	Business being empowered through data marketplace or data catalogue to automatically and easily search and request for data products, view business rules & metrics, and enable data product delivery. The tool provides insights into user ratings, orders placed, marketplace/catalogue usage, user profiles, subscriptions, user activity, access issues, recommendations, and tracks orders.

Perspectives: People, Process, Technology Reference: B, M, O, Z

C4. Production & Sharing

Active number of data products in production, exchanged across domains, and reused to serve business use cases to generate value.

Question: *What is the level of data product production and cross-domain interconnectivity within your organisation which generates business value?*

Level 0	Level 1	Level 2	Level 3
No or low number of data products in production. Data as an asset to be collected. Data products are designed for immediate use only. No data product re-usability across business functions.	Increased number of data products in production. However, cross-domain interconnectivity is still low. Data products are not serving use cases from other domains.	Increased number of data products in production. Data products are serving cross-domain use cases. Increased level of cross-domain interconnectivity.	High level of data products produced. Data as an asset to be shared. Data products are live and serve (cross-domain) multiple use cases, which reduce costs, promote innovation, and collaboration. High level of cross-domain interconnectivity.

Perspectives: People, Process Reference: I, K, L, M, N, Z

C5. Quality

Integrated data quality checks (accuracy, completeness, consistency, timeliness, uniqueness, and validity) to observe DQ-score to ensure data is delivered as a trustful product.

Question: *Is data quality tracked, monitored, and shared for the data products in your organisation?*

Level 0	Level 1	Level 2	Level 3
No data quality tracked, monitored, or shared for data products.	Data quality is only measured ad hoc for some data products. This is not automatically integrated. No clear DQ-issue protocol. No clear monitoring in place to track the data product quality.	Data quality is tracked and monitored. However, not completely an automated process or providing all relevant metrics to the data product consumers.	Automatically tracking, monitoring, and sharing data quality over all dimensions. Number of DQ-issues reported. DQ-score is provided by accessing data products in the discovery tool to ensure data is delivered as a trustful product. Next to objective measures, quality perception and experience are measured from the user's perspective.

Perspective: Process Reference: G, L, Z

C6. Ontology

Enabled graph consisting of the domains and the globally interoperable data products, which describe the relations and dependencies amongst the domains.

Question: *Is the ontology in your organisation available as an overview of all the relations and dependencies amongst the domains based on their shared and derived data products?*

Level 0	Level 1	Level 2	Level 3
Ontology not represented. No understanding of the interdependencies and interconnectivity of domains.	Ontology is not explicitly available. However, there is an understanding amongst the data product producers and consumers of how data is shared across the domains.	Ontology available, but not automatically updated while data is requested and shared.	Business domains and their relationships are represented in a graph to obtain real-time insights into all interoperable data products, domain interconnectivity, and domain dependencies. The ontology is automatically updated.

Perspectives: People, Process Reference: I, M, Z

C7. Archetypes

Categorised the data products into the three archetypes of domain-oriented analytical data: sourced-aligned, aggregated, and consumer-aligned, to understand data product usage, address data quality issues more easily, and improving business needs alignment.

Question: *Are the data products categorised into the three archetypes of domain-oriented analytical data?*

Level 0	Level 1	Level 2	Level 3
Data products are not categorised. No distinction is made between the different archetypes. No understanding of the differences between the archetypes.	Awareness about the differences of data product archetypes. However, the categories are not documented and discoverable for data product consumers.	The archetypes are provided and discoverable for almost all data products in the discovery tool. Adding the archetype category has been done manually. Guidelines are defined whether it is a source-aligned, aggregated, or consumer-aligned data product.	All data products have been categorised according to the three archetypes by integrated processes or standards during product development. Archetypes could be searched for in the data discovery tool. Categorisation helps understand data product usage, addresses data quality issues more easily, and improves business needs alignment.

Perspective: Process Reference: Z

C8. Structural Components

Embedded types of structural components (code, data, metadata, policy, and specifications of infrastructure dependencies) in the data products.

Question: *Are the data products consisting of the different structural components?*

Level 0	Level 1	Level 2	Level 3
Data is shared without its structural components. Data products may be published without a complete set of metadata. Data as a by-product of code. No policies embedded in the data product.	Some data products have the structural components included. However, this is not consistently applied to all data products. The structural components are not embedded in the data product development process.	Almost all structural components are consistently included for the data products by the different domains. Including the components is not yet completely integrated into the data product development process. Still, most of the data products consist of all components.	All data products consist of all embedded types of structural components, due to integration in the development process. Data and code as one unit. All data products require business, technical, and operational metadata. Embedded policies are validated and imposed in the flow of the data product's life-cycle, which allows continue testing and enforcing policies.

Perspective: Process Reference: Z

C9. Lead Time

Lead time to successfully build, test, deploy, discover, use, or change data products or its policies.

Question: *What is the current state of the overall lead time for building, testing, deploying, discovering, using, or changing data products or its policies in your organisation?*

Level 0	Level 1	Level 2	Level 3
Relatively high overall lead time for building, testing, deploying, discovering, using, or changing data products or policies. All processes are still manually and there is a high dependency on the central IT or Analytics team to perform tasks, and get access or approvals.	High overall lead time for building, testing, deploying, discovering, using, or changing data products or policies. However, processes are getting more automated. Central IT or Analytics team is still performing most tasks and providing access or approvals.	Reduced overall lead time to successfully build, test, deploy, discover, use, or change data products or policies. Increased level of domain autonomy, due to improved automated processes and training. Domains operate more autonomously and are less dependent on central IT or Analytics team.	Relatively low overall lead time to successfully build, test, deploy, discover, use, or change data products or policies. Processes are completely automated. Domains operate autonomously and are not or only less dependent on central IT or Analytics team.

Perspective: Process Reference: Z

C10. Discoverability

Enabled ability for data users to search, explore, and request the available data products.

Question: *Are data products discoverable in your organisation?*

Level 0	Level 1	Level 2	Level 3
No ability to search, explore, and request data products.	Enabled possibility to search, explore, and request data products by data discovery tool. Not all data products are stored. It is a difficult and unclear process. Overall discovery lead time is therefore high.	Data products could be explored, searched, and found in the available data discovery tool. All data products are stored in the tool. The overall discovery lead time is still relatively high.	Users could easily explore, search, and find available data products. Advanced searchability is implemented through search. Data products overview in catalogue or marketplace. Ontology explorer and discovery tool powers data product discovery at a global scale. Relatively low discovery lead time.

Perspectives: People, Process Reference: K, M, Z

C11. Addressability

Offered permanent and unique address to programmatically or manually access data products. Provided aggregate root (entry to all information about a data product, including documentation, SLO, and the data it serves) and schema evolution (if data is changed, versioning and communication of change needed).

Question: *Are data products addressable in your organisation?*

Level 0	Level 1	Level 2	Level 3
Data products are not addressable. A permanent and unique address is not available for data products.	Understanding of the importance of providing a permanent and unique address. This is provided for some data products, but not consistently for all data products. Aggregate root and schema evolution are missing.	Provided permanent and unique address for data products. Aggregate root and schema evolution are available, but there may be variability in their quality or completeness across different data products.	Permanent and unique address provided for programmatically or manually accessing data products. Standardised automated process to request access using workflow in place. All data products have a correct addressable aggregate root that serves as an entry to all documentation about a data product. Schema evolution is correctly included for versioning.

Perspective: Process Reference: K, M, Z

C12. Trustworthiness

Secured trust and confidence in data products by communicating SLO, data quality, lineage, approved usages by domains or individuals, metadata, code, and consumer experience or user satisfaction (reviews, star-rating, and net promoter score).

Question: *Are data products trustful in your organisation?*

Level 0	Level 1	Level 2	Level 3
No trust in data products. No effort to communicate SLO, data quality, lineage, approved usages by domains or individuals, metadata, code, and consumer experience.	Remained low trust in data products. SLO, data quality, lineage, approved usages by domains or individuals, metadata, code, and consumer experience are ad-hoc (not) included. These components are also often lacking consistency or are incomplete.	Increased trust in data product. SLO, data quality, lineage, approved usages by domains or individuals, metadata, code, and consumer experience are mostly included and are complete.	There is high trust and confidence in data products. This is obtained by communicating the SLO, data quality, lineage, governance assured by processes and platform, approved usages by domains or individuals, metadata, code, and consumer experience or user satisfaction (reviews, star rating, and/or net promoter score) which are all of high quality.

Perspectives: People, Process Reference: K, M, Z

C13. Descriptiveness

Provision of meaning to understand data product, by including title, description, tags, data governance markings, and manual.

Question: *Are data products described in your organisation?*

Level 0	Level 1	Level 2	Level 3
No descriptions provided for the data products. It is hard to understand data products.	Some data products have provisioning of meaning, e.g., by titles, descriptions, tags, or manuals. Overall, it is still difficult to understand the data products.	Most of the data products have provisioning of meaning, e.g., by titles, descriptions, tags, or manuals. Understanding the data products has become easier.	All data products have clear provision of meaning, by including title, description, tags, data governance markings, and user manual. The lineage of data provenance is documented. The ontology manager enriches further metadata on objects and their properties.

Perspective: Process Reference: K, M, Z

C14. Interoperability

Enabled ability to correlate data products across domains and stitch them together (join, filter, aggregate).

Question: *Are data products interoperable in your organisation?*

Level 0	Level 1	Level 2	Level 3
No or low data product interoperability. It is not possible or hard to derive new data products from current data products.	Data product interoperability occurs. Due to a lot of data products not consistently having standardised field types, metadata, or polysemes identifiers, it is hard to correlate data products.	Increased data product interoperability. Most of the data products have for example common field types, polysemes identifiers, unique addresses, and metadata. Increased number of derived data products coming from different domains.	High level of correlation across domains by derived data products. There is a high ability to correlate data products and stitch them together, enabled by standardised and provided field types, polysemes identifiers, data product global addresses, common metadata fields, schema linking, data linking, and schema stability.

Perspectives: People, Process Reference: K, M, Z

C15. Security

Secured data products which consist of policies (as code which could be versioned, automatically tested, deployed and observed, and computationally evaluated and enforced) addressing access control, encryption (by default), data classification, protection, confidentially levels, data retention, and other regulations.

Question: *Are data products secure in your organisation?*

Level 0	Level 1	Level 2	Level 3
Data products are not secure. No policies included for data protection.	Security level varies across the data products. Some data products have policies embedded related to access or usage whereas others are still insecure.	All or almost all security policies are consistently embedded in the data products. However, this process is not completely automated in the development process.	High level of data product security by automatically implementing policies (as code which could be versioned, automatically tested, deployed and observed, and computationally evaluated and enforced) which address access control, encryption (by default), data classification, protection, confidentially levels, data retention, and other regulations.

Perspective: Process Reference: K, M, Z

C16. Accessibility

Enabled possibility for various users to access and shop data products.

Question: *Are data products accessible in your organisation?*

Level 0	Level 1	Level 2	Level 3
No possibility for business units to get access to data products. Only the IT or Analytics team have access to data products.	Difficult process or high requirements set for getting access to data products by the business. Approvals are only provided in cases of high priority.	Increased accessibility to data products. However, access permission is still limited to a certain amount of people across the domains.	All users across the domains have the possibility to access and shop data products easily.

Perspectives: People, Process Reference: K, M, Z

Dimension D: Self-Serve Data Infrastructure as a Platform

D1. Infrastructure & Platform

Established infrastructure as a platform which provides the set of technologies to enable domain teams to search and share data products, to create and set policies for data products, and for storing, computing, and caching purposes.

Question: *Is the infrastructure in place to search and share data products, create and set policies, and for storing, computing, and caching information?*

Level 0	Level 1	Level 2	Level 3
No infrastructure in place. There are no technologies for sharing or storing data products.	Initial set of technologies available, but do not function as one coherent platform. The extent to which the infrastructure is implemented differs heavily across the domains.	Set of technologies is established. Almost all domain teams are enabled to search and share data products. However, not all technologies provide automated processes.	Fully established deployment and configuration of infrastructure as a coherent platform. The platform provides a set of automated technologies to enable domain teams to search and share data products, create and set policies for data products, and store, compute, and cache information.

Perspectives: People, Process, Technology Reference: C, G, I, Z

D2. Self-Serve

The extent to which domains are autonomously able to use the underlying data infrastructure capabilities by self-serve API's for creating, sharing, storing, computing, processing, and querying tasks.

Question: *Are the domains in your organisation able to autonomously perform all the activities in a self-serve manner?*

Level 0	Level 1	Level 2	Level 3
Platform is not self-serve. Full service of data activities by the central IT or Analytics team.	Initial self-serve data platform. Domains are enabled to perform some tasks by themselves. Other tasks still have to be performed by the central IT or Analytics team.	Improved self-serve data platform. Domains are enabled to do most of their tasks autonomously by the established technologies. Only a few functionalities are not self-serve.	Everything on the platform is designed to be self-serve. Domains are enabled by the self-serve platform to autonomously perform creating, sharing, storing, computing, processing, and querying tasks.

Perspectives: People, Process, Technology Reference: Z

D3. Performance

Indicated effectiveness and efficiency of the platform, in terms of platform costs, effort per environment request, user experience, number of people using the platform, number of data products using the platform, level of automation, number of features available, etc.

Question: *How would you describe the performance of the platform in your organisation?*

Level 0	Level 1	Level 2	Level 3
Poor performance of infrastructure. Relatively high platform costs, high effort per environment request, negative user experience, low number of people using the platform, low number of data products using the platform, no automation, and low number of features available.	Emerging infrastructure performance. Reduced platform costs, reduced effort per environment request, improved user experience, increased number of platform participants, increased number of data products available, more processes automated, and increased number of platform features available.	Consistent infrastructure performance. Performance is meeting its expectations in terms of platform costs, effort per environment request, user experience, number of platform users, number of data products available, level of automation, and number of platform features.	High effectiveness and efficiency of the platform. Relatively low platform costs, low effort per environment request, highly positive user experience, a high number of people using the platform, a high number of data products using the platform, completely automated, and a high number of features available.

Perspectives: People, Process, Technology Reference: Z

D4. Ownership

Establishment of the data platform product ownership role, defined data platform owner, and taking ownership and responsibility to deliver the platform services as a product.

Question: *Are the data platforms having a data platform product owner who takes the data platform responsibilities?*

Level 0	Level 1	Level 2	Level 3
No data platform product owner defined. All responsibilities are fragmented or non-allocated.	Data platform product owner defined and allocated. However, it still needs to be determined what the exact responsibilities are and how they will be measured.	Data platform product owner defined and allocated. Responsibilities determined and performed. However, performance is not measured.	Establishment of the data platform ownership role, allocated data platform owner, and taking ownership and responsibility. The owner creates a successful ecosystem of data infrastructure to deliver the platform services as a product. Key metrics to track performance are measured and accomplished, e.g., a robust number of data products in production or satisfaction score.

Perspectives: People, Process Reference: Z

D5. Platform Team

Composed platform team managing, building, and maintaining the self-serve data platform.

Question: *Is there a central team managing, building, and maintaining the self-serve data platform?*

Level 0	Level 1	Level 2	Level 3
No central platform team defined. Responsibilities are fragmented to individuals or non-allocated.	Initial central platform team defined and composed. However, number of representatives and responsibilities is limited. Platform improvements are ad-hoc initiated by the team.	Central platform team defined and composed. Increased number of representatives and responsibilities. Improvements are integrated more frequently, but still ad-hoc.	Central platform team defined and composed. The platform is built, maintained, managed and created as self-serve by a central dedicated platform team. The platform team continuously seek to find new and improved ways to utilise automation, remove friction in data sharing, and optimise the user experience.

Perspectives: People, Process, Technology Reference: Z

D6. Multiplane Platform

Established multiplane data platform with declarative interfaces: mesh experience plane, data product experience plane, and infrastructure utility plane.

Question: *Are the mesh experience plane, data product experience plane, and infrastructure utility plane implemented as distinctive planes in your organisation to distinguish between different classes of platform services?*

Level 0	Level 1	Level 2	Level 3
Separation of planes is not respected.	Some functionalities of the different planes are separated. However, no clear distinction has been established in general.	The different planes are separated to some extent. However, no complete distinction is established. Decisions made by a single plane could still impair the other planes.	The mesh experience plane, data product experience plane, and infrastructure utility plane are implemented as distinctive planes. The different planes are optimised and specific single-plane decisions do not impair the developments on other planes.

Perspectives: Process, Technology Reference: Z

D7. Analytical API's

Created analytical data sharing interfaces to domains to enable autonomy. Analytical interfaces are the API's that data products expose to get discovered, understood, observed, and to share the data products.

Question: *Are the domains in your organisation having analytical data sharing interfaces?*

Level 0	Level 1	Level 2	Level 3
Domains do not have analytical data sharing interfaces.	Only few domains have analytical data sharing interfaces. Other domains are not able to discover, understand, observe, or share data products.	Almost all domains have analytical data sharing interfaces. Only a few are not able to discover, understand, observe, or share data products.	Enabled autonomy for all domains by analytical data sharing interfaces. Domains enable data products expose to get discovered, understood, observed, and to share the data products.

Perspectives: People, Process, Technology Reference: Z

D8. Operational API's

Created operational interfaces to domains to enable autonomy. Operational interfaces are the API's and applications through which a business domain shares its transactional capabilities and states with the wider organisation.

Question: *Are the domains in your organisation having operational data sharing interfaces?*

Level 0	Level 1	Level 2	Level 3
Domain do not have operational data sharing interfaces.	Only few domains have operational data sharing interfaces. Other domains are not able to discover, understand, observe, or share their capabilities.	Almost all domains have operational data sharing interfaces. Only a few are not able to discover, understand, observe, or share their capabilities.	Enabled autonomy for all domains by operational interfaces to share its transactional capabilities and state with the wider organisation.

Perspectives: People, Process, Technology Reference: Z

Dimension E: Federated Computational Governance

E1. Security & Compliance

Defined security policies (data usage, access approval, retention, archival, or GDPR regulations) and having security governance tools in place to check in a self-serve manner workflows automatically. Alerting and notifying in case of risk or violating the policies.

Question: *Are security policies defined and are tools established for monitoring compliance of all regulations and agreements?*

Level 0	Level 1	Level 2	Level 3
No security policies defined. No tools established to monitor the policies. No understanding of data security and compliance.	Initial set of security policies defined. No or only a few tools in place which are used to check security and compliance. The monitoring process is manual. Tools are only used in case an incident happened.	Extensive set of security policies defined. Governance tools are in place for monitoring compliance and are frequently checked. However, workflows still need to be checked manually.	Defined all security policies (data usage, access approval, retention, archival, or GDPR regulations) and security governance tools are in place in a self-serve manner to check workflows automatically. Alerting and notifying in case of risk or violating the policies.

Perspectives: Process, Technology Reference: C, I, K, M, Z

E2. Global Policies

Defined global communication protocols and standards that cover the full scope of central governance. Protocols govern how data products express their semantics, format, query language, and what SLO each guarantee.

Question: *Are global communication protocols and standards defined to enable interoperability?*

Level 0	Level 1	Level 2	Level 3
No global protocols and standards defined.	Some initial global communication protocols and standards defined. Only a few essential security, privacy, and standardisation considerations are covered. Interoperability issues still occur frequently.	Extensive set of global communication protocols and standards defined. However, not all essential security, privacy, and standardisation considerations are covered. Hardly any interoperability issues.	Defined global communication protocols and standards that cover the full scope of central governance. Protocols in place govern how data products express their semantics, format, query-language, and what SLO each guarantee. Global policies have fully enabled interoperability. All essential security, privacy, and standardisation considerations are covered.

Perspective: Process Reference: O, Z

E3. Federated Policies

Adopted federated policies and standards by domains that guide the decision-making process and decide about what policies the organisation must implement globally and what could be left to domains.

Question: *To what extent have domains adopted the federated governance operation?*

Level 0	Level 1	Level 2	Level 3
No domains have adopted the federated governance operation. Data governance from a top-down centralised operational model. Data is still governed by a central function.	Only few domains have adopted the federated governance operation. Most of the domains still act according to global top-down policies.	Almost all domains have adopted the federated governance operation. The federated approach is almost enabled.	All domains have adopted the federated governance operation. On the local level is decided what policies the organisation must implement globally and what could be left to domains. The full federated approach is enabled.

Perspective: Process Reference: O, Z

E4. Monitoring

Automated continuously monitoring of policy enforcement. E.g., monitor the operational health of the mesh (check data expectations by statistical properties, check transformations, serve data in compliance with SLO), debug and perform post-mortem analyses, perform audits (regulatory, risk, security), and understand data lineage (data origin and transformation tracking).

Question: *Are the data governance policies continuously and automatically monitored?*

Level 0	Level 1	Level 2	Level 3
No policies are continuously and automatically monitored. Monitoring not in place, or still completely manual.	Only a small set of policies are continuously and automatically monitored. If monitoring tooling is in place, the process is still manual.	Extensive set of policies are continuously and automatically monitored. Only a few monitoring tasks were performed manually.	All policies are automated and continuously monitored. E.g., the operational health of the mesh (check data expectations by statistical properties, check transformations, serve data in compliance with SLO), debug and perform post-mortem analyses, perform audits (regulatory, risk, security), and understand data lineage (data origin and transformation tracking).

Perspectives: Process, Technology Reference: Z

E5. Standardisation

Enforced global standards to enable interoperability. E.g., consistent processes and tools; global definitions for domains, products, and ownership in the glossary; data product business and technical KPI's; entity ID's (polysemes identifiers); data sharing API's; common metadata fields including SLO's, documentation, and modelling language; schema linking; data linking; schema stability (backward compatibility).

Question: *To what extent is standardisation integrated and helps it to enable interoperability?*

Level 0	Level 1	Level 2	Level 3
No standardisation enforced to enable interoperability. Not possible to correlate or share data products across domains.	Low level of standardisation. Difficult to enable interoperability, due to lack in amongst others consistency across global processes and tools, definitions, and KPI's.	High level of standardisation. Interoperability enabled by amongst others consistency across global processes and tools, definitions, and common metadata fields including SLO's and KPI's.	Standardisation enforced wherever needed, to fully enable interoperability. Consistent processes and tools in place, global definitions in glossary, data product business and technical KPI's, entity ID's, data sharing API's, common metadata fields including SLO's, documentation and modelling language, schema linking, data linking, and schema stability.

Perspectives: Process, Technology Reference: Z

E6. Computational Policies & Automation

Codified and automated governance policies. Computational governance standards as code, policies as code, and automated tests. Policies as automated and machine-led processes embedded in each and every domain and data product.

Question: *To what extent are the governance policies and standards codified and automated?*

Level 0	Level 1	Level 2	Level 3
No computational policies and automation in each and every domain and data product. Processes primarily with human intervention.	Only few domains have automated processes embedded. Some data products have standards as code and embedded automated tests.	Most of the domains have automated processes embedded. Almost all data products have standards as code and embedded automated tests.	Policies as automated and machine-led processes embedded in each and every domain. All data products on the mesh are embedding codified and automated governance policies. Policies and standards as code and automated tests.

Perspectives: Process, Technology Reference: N, O, Z

E7. Governance Team

Allocated central governance team consisting of domain, platform, and SME (legal, compliance, security) representatives, which define the central policies for the platform to build in, which could be run, and monitored automatically.

Question: *Is a central governance team allocated to define the central policies which need to be integrated into the platform?*

Level 0	Level 1	Level 2	Level 3
No central governance team allocated to define central policies.	Initial central governance team defined and composed. However, number of representatives and responsibilities is limited. Governance improvements are ad-hoc initiated by the team.	Central governance team defined and composed. Increased number of representatives and responsibilities. Improvements are integrated more frequently, but still ad-hoc.	Central governance team is allocated. The team consists of all domain, platform, and SME (legal, compliance, security) representatives. Central policies are defined, updated, improved, and monitored frequently.

Perspectives: People, Process Reference: Z

E8. Incident Management

Integrated incident management to detect root cause issues and recover through platform's automated processing.

Question: *Are there incident management capabilities integrated to detect errors and recover through automated processing?*

Level 0	Level 1	Level 2	Level 3
No incident management capabilities considered or implemented. Issues are ignored.	Some initial incident capabilities established. Incidents are often addressed reactively on an ad-hoc basis without formal processes.	Automated tools and processes in place to detect errors. However, these incident management capabilities are not able to recover from issues automatically.	Incident management capabilities completely integrated. Capabilities detect root cause issues and recover through automated processing.

Perspectives: Process, Technology Reference: Z

6 Model Outcomes

Chapter 5 motivated the design choices and eventually presented and defined the developed DMMAM-elements. The next step will look into which results could be obtained once the assessment has been filled in. Chapter 6 therefore aims to provide the answer to the question of what outcomes could be provided by using the DMMAM. The sub-sub-question, sub-sub-objective, and sub-sub-deliverable for this chapter are provided below.

Sub-Sub-Question 2.3

What outcomes could be provided by using the data mesh maturity assessment model?

Sub-Sub-Objective 2.3

Defining the outcomes of the data mesh maturity assessment model.

Sub-Sub-Deliverable 2.3

Outcomes of the data mesh maturity assessment model.

The structure of this chapter follows in total five sections. First of all, the most important maturity score metrics will be introduced and defined in Section 6.1. The same metrics will be approached from PPT-perspective in Section 6.2. By default, the maturity scores are calculated by equally weighting the dimensions and characteristics. However, Section 6.3 presents the possibility for experimentation with unequal weights. Section 6.4 presents charts which are useful for providing insights visually. Lastly, the insights presented in Section 6.1, 6.2, 6.3, and 6.4 are obtained from an individual self-assessment or from assessments conducted by a single organisation. Section 6.5 presents the benchmarking option which enables the comparison of maturity scores across multiple organisations.

Before presenting the metrics, an explanation of the self-assessment will be provided. It has been decided to employ an Excel-based questionnaire as the assessment tool for the DMMAM. This choice is particularly appropriate argues Lindemulder (2015) since this method is used by various, different, but comparable maturity assessment models.

Appendix F presents a manual of the various sheets of the designed Excel-based DMMAM. These instructions show what the models look like and provide descriptions of all the steps, explaining the capabilities of the model. Furthermore, Section 12.2.1 will delve into deployment suggestions from the client's perspective, showcasing how the model could be utilised to serve users prior to, during, and after conducting an assessment. At the time of self-assessment, participants will only receive the first and second Excel-sheets. The self-assessment guides the participant through questions about all characteristics. The participant needs to assess for each characteristic of where the organisation is currently positioned on the maturity scale. In addition, participants are also asked to indicate a target maturity state for a predetermined scope. The answers *Level 0*, *Level 1*, *Level 2*, and *Level 3* are to be selected from a drop-down menu. Additionally, the possibility to select the answer *Unknown* is added. In case the participant does not provide any answer to the question, the field remains blank. In conclusion, six different answers are possible which all would impact the final maturity scores. Lastly, based on the suggestion from Participant D, as presented in Table B.19, a comment section is added next to each characteristic, to offer the participant the possibility to add notes, remarks, and questions. Appendix F, manual page 3 and 4, provide how it looks for the participants.

6.1. Metrics

After having completed the self-assessment, various scores will be calculated automatically in Excel. The main metric is the overall maturity score which covers all dimensions and dimensional-specific maturity scores. Additionally, including the unknown rate and response rate helps understand whether participants were unable to answer or chose not to answer the question, including those who may have forgotten to respond. Table 6.1 shows the metrics and definitions established by the designer.

Table 6.1: Maturity Metrics

Nr.	*Metric*	*Definition [Unit]*
1	**Current Maturity Level**	Average current maturity level for all characteristics belonging to all/the specific dimension(s). [Rational Number]
	Unknown Rate	The number of questions where the current maturity level is answered by *Unknown* relative to the total number of questions where the current maturity level is provided. [Percentage]
	Response Rate	Number of questions where the current maturity level is provided relative to the total number of questions. [Percentage]
2	**Target Maturity Level**	Average target maturity level for all characteristics belonging to all/the specific dimension(s). [Rational Number]
	Unknown Rate	The number of questions where the target maturity level is answered by *Unknown* relative to the total number of questions where the target maturity level is provided. [Percentage]
	Response Rate	Number of questions where the target maturity level is provided relative to the total number of questions. [Percentage]
3	**Maturity Gap**	The absolute difference between the current maturity level and the target maturity level for all/the specific dimension(s). [Rational Number]

6.2. People, Process, Technology Perspectives

Instead of looking into the scores for the different dimensions, the metrics as introduced in Table 6.1 will also be calculated for the PPT-perspectives to obtain an understanding of their progress accordingly. Table 6.2 shows the definitions as established by the designer.

Table 6.2: Maturity Metrics: People, Process, Technology

Nr.	*Metric*	*Definition [Unit]*
1	**Current Maturity Level**	Average current maturity level for all characteristics belonging to the specific perspective. [Rational Number]
	Unknown Rate	The number of questions where the current maturity level is answered by *Unknown* relative to the total number of questions where the current maturity level is provided. Only the questions related to the specific perspective are considered. [Percentage]
	Response Rate	Number of questions where the current maturity level is provided relative to the total number of questions. Only the questions related to the specific perspective are considered. [Percentage]
2	**Target Maturity Level**	Average target maturity level for all characteristics belonging to the specific perspective. [Rational Number]
	Unknown Rate	The number of questions where the target maturity level is answered by *Unknown* relative to the total number of questions where the target maturity level is provided. Only the questions related to the specific perspective are considered. [Percentage]
	Response Rate	Number of questions where the target maturity level is provided relative to the total number of questions. Only the questions related to the specific perspective are considered. [Percentage]
3	**Maturity Gap**	The absolute difference between the current maturity level and the target maturity level. Only the questions related to the specific perspective are considered. [Rational Number]

6.3. Experimentation

While calculating the maturity scores, as presented in Section 6.1 and 6.2, all characteristics and dimensions have equal weights by default. An optional function has been added to the assessment document in which it becomes possible to experiment with different weights. This option enables the assignment of critical and non-critical characteristics and dimensions. The assessment has the possibility to change both the strength of the weights and the distribution of weights over the characteristics and dimensions. As a result, weighted averages will be presented for the overall score and dimensional-specific scores. An overview of the metrics and definitions established by the designer is provided in Table 6.3.

Table 6.3: Maturity Metrics: Experimentation

Nr.	*Metric*	*Definition [Unit]*
1	Current Maturity Level (Dimension)	Weighted average current maturity level for all characteristics belonging to all dimensions. [Rational Number]
2	Target Maturity Level (Dimension)	Weighted average target maturity level for all characteristics belonging to all dimensions. [Rational Number]
3	Maturity Gap (Dimension)	The absolute difference between the weighted average current maturity level and the weighted average target maturity level for all dimensions. [Rational Number]
4	Current Maturity Level (Characteristics)	Weighted average current maturity level for all characteristics belonging to all/the specific dimension(s). [Rational Number]
5	Target Maturity Level (Characteristics)	Weighted average target maturity level for all characteristics belonging to all/the specific dimension(s). [Rational Number]
6	Maturity Gap (Characteristics)	The absolute difference between the weighted average current maturity level and the weighted average target maturity level for all/the specific dimension(s). [Rational Number]

6.4. Charts

The insights will be presented visually in a *bar chart* and *radar chart*. The bar chart will be added since it could easily rank the dimensions and characteristics from immature to mature. An example bar chart for dimension A, ranked on target maturity level, is provided in Figure 6.1. In addition, the radar chart visualises the maturity gap by presenting both the current and target maturity states. An example radar chart for dimension A is provided in Figure 6.2.

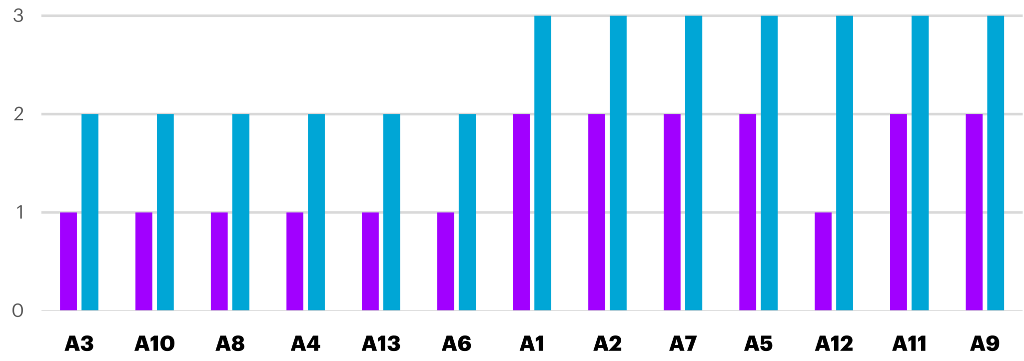


Figure 6.1: Bar Chart

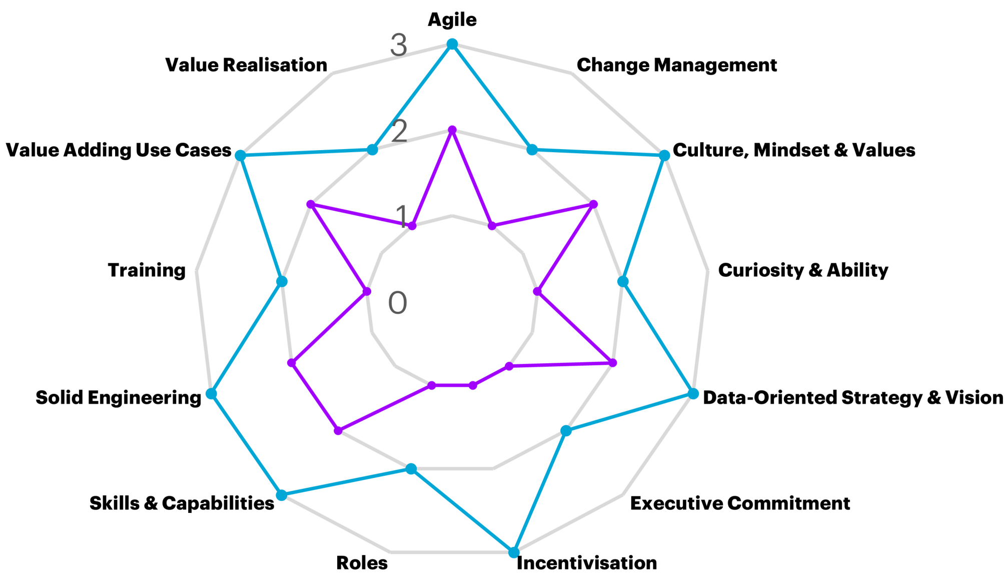


Figure 6.2: Radar Chart

6.5. Benchmarking

Benchmarking is enabled by comparing the metrics over multiple organisations by the client. This implies that the organisation could only be compared to other organisations where the client has previously conducted an assessment. Therefore, it does not pertain to a publicly accessible database. The current maturity score, target maturity score, and maturity gap could be presented for all dimensions and perspectives. In addition, a bar chart could be added to present the current and target maturity states over the dimensions across organisations. At last, radar charts could present the current maturity levels per dimension and the target maturity level per dimension over the organisations.

Conclusion Phase C

Phase C aimed to design and develop the DMMAM. This section concludes what model has been designed to assess the maturity of a data mesh implementation within an organisation.

The systematic engineering design approach from Dym et al. (2013) was used to provide structure to the design process. This approach consists of objectives, constraints, and functional analyses, aiming for selecting a preferred model design from the design space. The objectives analyses showed that in aiming for a successful DMMAM, balance is needed between having a *valuable* model and a *feasible* assessment. This trade-off between obtaining useful results while dealing with limited resources was also reflected in all constraints. This implied the number, exclusiveness, and exhaustiveness of the data mesh characteristics; the number of maturity levels; the number and background of participants in the assessment; and the duration of assessment discussion sessions. The preferred design showed that the model's functions will evaluate the maturity of a data mesh in an overall score, individual dimensional scores, and scores from PPT-perspectives. In addition, guidance will be provided by comparing the current and target maturity state as a maturity gap and by enabling the possibility of benchmarking multiple organisations.

By establishing means for the various functions and adhering to the need to have a self-assessment, the DMMAM developed in this research consists of the following six model elements: *maturity levels*, *dimensions*, *characteristics*, *perspectives*, *questions*, and *criteria and requirements*. The maturity will be evaluated by four maturity levels, which are classified as Level 0: *Non-Initiated*, Level 1: *Conceptual*, Level 2: *Defined*, and Level 3: *Achieved*. Calculating the scores is performed by considering equal weights for the elements. Data mesh will be evaluated by five dimensions as components, which are A. *Data Foundation & Organisational Change*, B. *Domain Oriented Decentralised Data Ownership & Architecture*, C. *Data as a Product*, D. *Self-Serve Data Infrastructure as a Platform*, and E. *Federated Computational Governance*. Dimension A refers to the generic data management, people, and cultural aspects which provide the foundation to start implementing the four other dimensions, as introduced by Dehghani (2022a). The five dimensions are elaborated by a total of 54 characteristics. Figure 6.3 presents the considered DMMAM dimensions, characteristics, and levels. For each of these characteristics, PPT-labels and a question are added. The self-assessment is presented as a multiple-choice questionnaire in which descriptions, in terms of criteria and requirements, are provided to help position the maturity of a characteristic at a particular level.

During the assessment, the participants will be asked to individually rate the 54 questions on the current and target level of maturity that are most fitting for their organisation. Each statement needs to be rated on a scale from Level 0: *Non-Initiated* to Level 3: *Achieved*, or as *Unknown*. Various results will be presented after conducting the assessment. As mentioned, an overall score, individual dimensional scores, and scores from PPT-perspectives will be calculated. In addition, the maturity gaps between the current and target maturity states will be determined. Moreover, an optional experimentation function has been added. This option enables the possibility to experiment with different weights and weight strengths to evaluate the impact of critical and non-critical elements. To obtain the insights visually, bar charts and radar charts are added to enable the possibility to rank the dimensions and characteristics and present the maturity gaps respectively. At last, the scores and charts will be used for benchmarking multiple organisations.

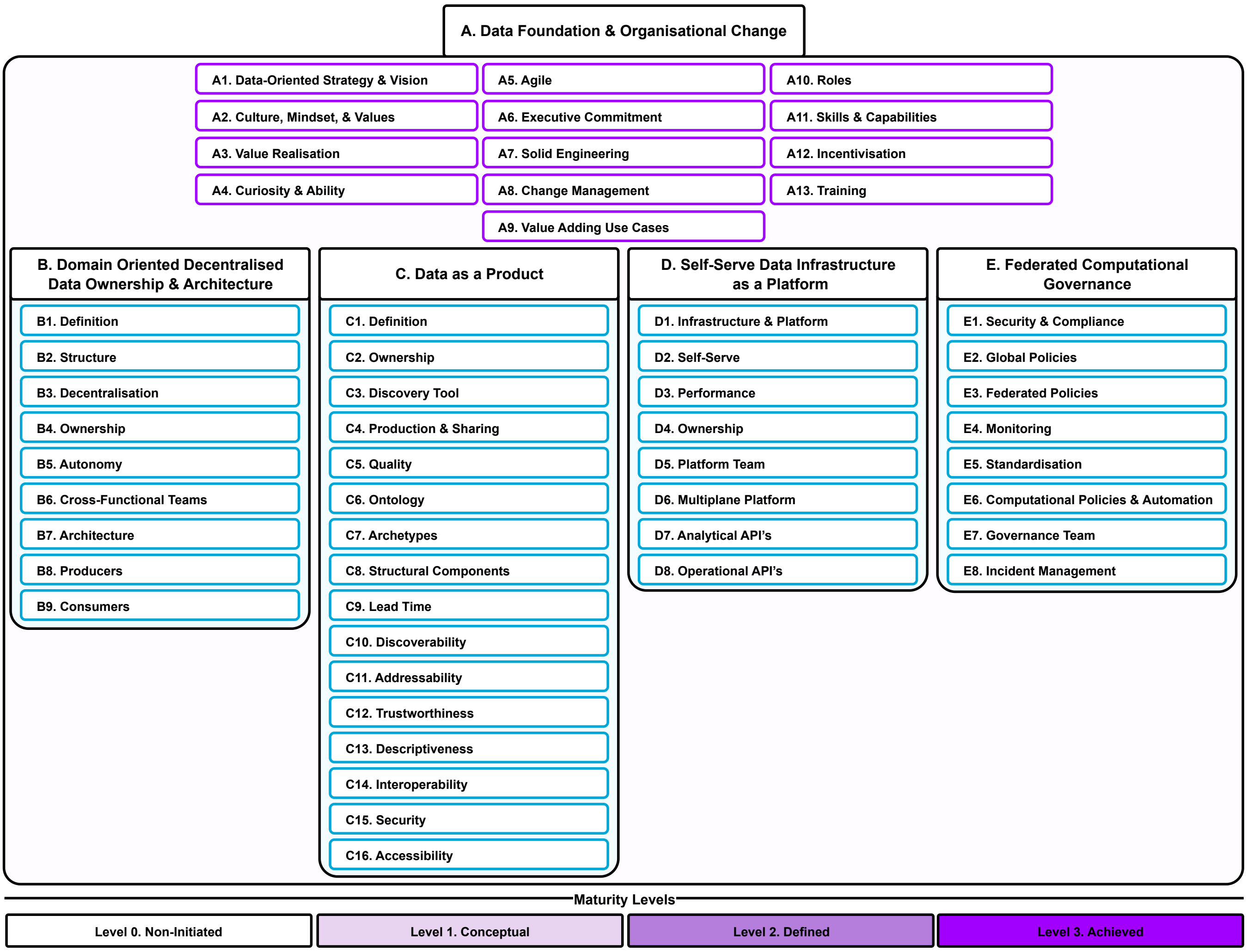


Figure 6.3: Model Dimensions, Characteristics, and Levels

Phase D

Demonstration

7

Model Demonstration

Phase D focuses on the demonstration of the DMMAM. Conducting assessments by using the designed DMMAM for multiple organisations in cases will demonstrate its applicability and usefulness. While applicability refers to the practical feasibility of demonstrating the DMMAM, usefulness pertains to the ability to provide the desired outcomes for the proposed maturity metrics and charts from Phase C. The decision to consider the applicability and usability is based on work from Gökalp et al. (2022), where these were used for demonstration and evaluation of their data maturity assessment model. The sub-question, sub-objective, and sub-deliverable of Phase D are provided below.

Sub-Question 3

How to demonstrate the data mesh maturity assessment model?

Sub-Objective 3

Observing the applicability and usefulness of the data mesh maturity assessment model.

Sub-Deliverable 3

Demonstrated data mesh maturity assessment model.

This chapter is structured in four sections. Section 7.1 presents the process of selecting organisations for participation in the cases. The case structure, containing the maturity assessment activities, will be described in Section 7.2. The outcomes of the assessments will be presented in Section 7.3. Findings from the individual assessments will be compared in Section 7.4. Eventually, Phase D will conclude regarding the demonstration of the DMMAM.

7.1. Organisation Selection

For conducting maturity assessments in cases, organisations will be involved. Section 5.2 presented in Table 5.4 next to the *model* constraints also *assessment* constraints which need to be adhered to in this research. Therefore, specific criteria have been established that both the organisations and the participants representing them must meet while selecting organisations. The first constraint referred to the number of participants in the case. Participant A argued that at least three participants must be involved. This constraint aims to balance the validity of the final maturity score and research feasibility. To elaborate on this, the fewer participants are involved, the greater the probability that a few insights will have a large impact on the final score (Holland et al., 1986). Additionally, there is also a chance that certain aspects of data mesh may go unaddressed because the limited knowledge of these few participants may prevent them from adequately answering the questions. Secondly, the group of participants needs to be balanced to guarantee sufficient knowledge is available for addressing all different perspectives on data mesh (Participants B; C; D; E; F; G; N). This implies that organisations which neither have started exploring nor implementing data mesh will not be considered for participation. Organisations meeting the two constraints fit for participation in the cases.

The functional analysis in Section 5.3 presented that the model design aims to provide guidance for achieving higher levels of maturity by adding a benchmarking functionality. Comparing the maturity scores with industry competitors is helpful to obtain an understanding of how an organisation is positioned within its industry. To enable this comparison, selecting organisations within the same industry is preferred over cross-industry comparison. Moreover, including organisations from the same industry would match the *most similar* method as the case selection procedure introduced by Seawright & Gerring (2008). In terms of representativeness, Seawright & Gerring mention that most similar cases will provide a strong basis for making generalisations. However, this study will only perform a few assessments, thereby precluding any empirically significant generalisations (Ferguson, 2004; Webb & Shavelson, 2005). However, having organisations from the same industry is not a strict requirement since the individual cases and cross-industry comparisons still help for demonstration. Moreover, conducting cases in different sectors could aid in observing the applicability and usability of the model under different circumstances (Gökalp & Martinez, 2021).

Three organisations were identified and engaged through personal outreach facilitated by Accenture's professional network. The three organisations are all large multinational businesses that lead in their respective industries. Table 7.1, 7.2, and 7.3 present, in addition to Table 5.1, the participants from Organisation I, II, and III respectively.

Table 7.1: Case I: Organisation, Representative, and Participants

	Organisation	
Nr	Industry	Number of Employees
I	Financial Services	20,000
	Representative	
ID	Role	Working experience
P	Data Officer	10 years
	Participant	
ID	Role	Working experience
P	Data Officer	10 years
W	Data Management Lead	10 years
X	Senior Data Officer	20 years

Organisation I is represented by three participants. The roles cover a *Data Officer*, *Data Management Lead*, and *Senior Data Officer*, having respectively 10, 10, and 20 years of working experience.

Table 7.2: Case II: Organisation, Representative, and Participants

Nr	***Organisation*** Industry	Number of Employees
II	Financial Services	40,000

ID	***Representative*** Role	Working experience
Q	Managing Business Architect	20 years

ID	***Participant*** Role	Working experience
Q	Managing Business Architect	20 years
Y	Data Analytics Product Manager	10 years
AA	Data Management Consultant	10 years

Organisation II is also represented by three participants. The roles cover a *Managing Business Architect*, *Data Analytics Product Manager*, and *Data Management Consultant*, having 10 to 20 years of working experience.

Table 7.3: Case III: Organisation, Representative, and Participants

Nr	***Organisation*** Industry	Number of Employees
III	High-Tech Manufacturing	40,000

ID	***Representative*** Role	Working experience
R	Chief Data Product Owner	10 years

ID	***Participant*** Role	Working experience
R	Chief Data Product Owner	10 years
AB	System Architect	15 years

Organisation III is represented by two participants. The roles cover a *Chief Data Product Owner* and *System Architect*, having respectively 10 and 15 years of working experience. Unlike Organisation I and II, Organisation III does not have at least 3 participants. However, it is expected this case could still produce meaningful results, but the fact only two participants were involved will be taken into account.

7.2. Case Structure

Having presented the organisations and participants involved in the cases, this section will continue by explaining what is considered by a case. The case reflects the process of conducting a data mesh maturity assessment by using the designed DMMAM. Section 3.1.3 elaborated on the assessment activities, as presented by DAMA International (2017). Section 3.1.3 described that the total assessment process consists of five consecutive steps, namely planning the assessment activities, performing the maturity assessment, interpreting the results, creating a targeted programme for improvements, and conducting re-assessments. The first three activities will be included in the cases. Creating a target programme for improvement and conducting a re-assessment will not be considered due to limited time and capacity. Furthermore, it has been decided to conduct the cases in an identical manner. The setup will be described below in consecutive steps according to the three assessment activities.

I. Planning Assessment Activities

1. The DMMAM will be explained in an introductory meeting with the user representative.
2. The questionnaire will be shared with the user representative.
3. The user representative will share the questionnaire internally with colleagues.

II. Performing Maturity Assessments

4. The user participants will conduct the self-assessment individually.
5. In case the participants need help, scheduling a discussion session with the designer is possible.
6. After conducting the self-assessments, the user representative will send back the completed assessment forms to the designer.

III. Interpreting Results

7. The designer will calculate on individual and aggregated levels the scores for the metrics as presented in Section 6.1 and will create charts as shown in Section 6.4.
8. The complete set of results will be sent back to the user representative.
9. The case will be completed by a final workshop. In this workshop, the results will be explained by the designer to all user participants. In addition, this session would open the possibility for the participants to ask questions and discuss the findings.

7.3. Assessment Outcomes

This section presents in Section 7.3.1, 7.3.2, and 7.3.3 the outcomes of the three individual assessments. The cases followed the structure as presented. However, no participants did request an individual discussion session in the end. The aggregated responses from each organisation according to all characteristics are presented in Appendix E.

7.3.1. Organisation I

Table 7.4 presents the aggregated outcomes for the individual dimensions and for all dimensions together, which will be explained. The current overall maturity level equals 0.48, which means that the implementation, understanding, and experience regarding data mesh are labelled as non-initiated to conceptual. Dimension E, *Federated Computational Governance*, has the highest dimensional maturity score, which equals 0.75. Dimension C, *Data as a Product*, shows a score of 0.08 the lowest maturity level. The difference between these two dimensions equals 0.67, which is too small to be considered noteworthy, as it does not vary more than one level in terms of maturity. Dimension C shows while having the lowest current maturity score, the highest unknown rate which equals 23%. Overall, the participants filled in all fields and therefore, the overall response rate equals 100%. The overall response rate for the target maturity levels equals 0%, due to the participants from the organisation were not able to assess the target maturity levels. As a result, no maturity gaps are provided. Participant P, W, and X indicated this had nothing to do with the DMMAM.

Table 7.4: Aggregated Outcomes Organisation I: Dimensions

Dimension	*Current Maturity Level*	*Unknown Rate*	*Response Rate*	*Target Maturity Level*	*Unknown Rate*	*Response Rate*	*Maturity Gap*
A	0.54	0 %	100 %	NA	0 %	0 %	NA
B	0.70	0 %	100 %	NA	0 %	0 %	NA
C	0.08	23 %	100 %	NA	0 %	0 %	NA
D	0.67	8 %	100 %	NA	0 %	0 %	NA
E	0.75	17 %	100 %	NA	0 %	0 %	NA
A-E	0.48	10 %	100 %	NA	0 %	0 %	NA

Table 7.5 provides the outcomes from the different perspectives for Organisation I. The outcomes according to the PPT-perspectives, which are 0.48, 0.44, and 0.63 respectively, show minimal differences.

Table 7.5: Aggregated Outcomes Organisation I: People, Process, Technology

Perspective	*Current Maturity Level*	*Unknown Rate*	*Response Rate*	*Target Maturity Level*	*Unknown Rate*	*Response Rate*	*Maturity Gap*
People	0.48	7 %	100 %	NA	0 %	0 %	NA
Process	0.44	12 %	100 %	NA	0 %	0 %	NA
Technology	0.63	10 %	100 %	NA	0 %	0 %	NA

Overall, the relatively low current maturity levels, the relatively high unknown rates, and the inability to fill in the target maturity states show that Organisation I is currently immature regarding data mesh.

7.3.2. Organisation II

Table 7.6 shows the aggregated outcomes for the individual dimensions and for all dimensions together. The current overall maturity level equals 1.36, which means that the implementation, understanding, and experience regarding data mesh are labelled as conceptual. Dimension B, *Domain Oriented Decentralised Data Ownership & Architecture*, shows the highest dimensional maturity score, which equals 1.63. Dimension C, *Data as a Product*, reflects a score of 1.02 the lowest maturity level. The difference between these two dimensions is 0.61. The unknown rates show that the participants were able to provide answers to almost all questions. Dimension C, which has the lowest maturity score, has the highest unknown rate which equals 2%. The response rate shows a 100% result. The target overall maturity level equals 2.97, which means that the organisation aims to achieve to successfully establish, optimise, and continuously improve the data mesh characteristics while having a high level of understanding and experience with the data mesh application. The difference between the highest and lowest dimensional target maturity level equals 0.06, which is negligibly small. This means that all dimensions have the same target level, which equals approximately 3.00. Relatively high levels of unknown rates are presented for all dimensions, due to one of the participants was not able to assess the target states. The response rate shows also for the target maturity levels 100% results. The overall maturity gap equals 1.61, which is perceived as a large gap. The individual dimensions show an almost equal maturity gap, where the maximum difference between the maturity gaps equals 0.61.

Table 7.6: Aggregated Outcomes Organisation II: Dimensions

Dimension	*Current Maturity Level*	*Unknown Rate*	*Response Rate*	*Target Maturity Level*	*Unknown Rate*	*Response Rate*	*Maturity Gap*
A	1.46	0 %	100 %	3.00	36 %	100 %	1.54
B	1.63	0 %	100 %	2.94	37 %	100 %	1.31
C	1.02	2 %	100 %	2.94	35 %	100 %	1.92
D	1.50	0 %	100 %	3.00	33 %	100 %	1.50
E	1.46	0 %	100 %	3.00	33 %	100 %	1.54
A-E	1.36	1 %	100 %	2.97	35 %	100 %	1.61

For Organisation II, Table 7.7 provides the outcomes from the different perspectives. The results according to the PPT-perspectives, which are regarding the current maturity levels 1.44, 1.32, and 1.45 respectively, show minimal variation. This negligible difference also holds for the target maturity scores and maturity gaps.

Table 7.7: Aggregated Outcomes Organisation II: People, Process, Technology

Perspective	*Current Maturity Level*	*Unknown Rate*	*Response Rate*	*Target Maturity Level*	*Unknown Rate*	*Response Rate*	*Maturity Gap*
People	1.44	0 %	100 %	3.00	35 %	100 %	1.56
Process	1.32	1 %	100 %	2.97	35 %	100 %	1.65
Technology	1.45	0 %	100 %	3.00	36 %	100 %	1.55

Overall, the scores show that the organisation has some initial data mesh implementation, has an understanding of the concepts, and started experimenting. Furthermore, the organisation aims to improve its capabilities to the maximum maturity level. The large maturity gaps emphasise the discrepancy between the current as-is and the future to-be state. The organisation aims to make significant progress in advancing the development of data mesh.

7.3.3. Organisation III

Table 7.8 shows the aggregated outcomes for the individual dimensions and for all dimensions together for Organisation III. The current overall maturity level is 0.79, indicating that the implementation, understanding, and experience related to data mesh are classified as almost conceptual. Among the dimensions assessed, Dimension B, *Domain Oriented Decentralised Data Ownership & Architecture*, demonstrates the highest maturity score of 1.17. On the other hand, Dimension C, *Data as a Product*, exhibits the lowest maturity level with a score of 0.34. There exists a disparity of 0.83 between these two dimensions implying that it does not vary much, as it is still less than one level in terms of maturity. The data reveals that all questions were answered by a provided maturity level, as indicated by the 0% unknown rate. Moreover, the response rate achieved a perfect score of 100%. However, the overall response rate for the target maturity levels stands at 0% because the participants from the organisation were unable to assess these levels. Consequently, no maturity gaps could be calculated. Also Participant R and AB indicated that not filling in the target maturity states had nothing to do with the DMMAM-design.

Table 7.8: Aggregated Outcomes Organisation III: Dimensions

Dimension	*Current Maturity Level*	*Unknown Rate*	*Response Rate*	*Target Maturity Level*	*Unknown Rate*	*Response Rate*	*Maturity Gap*
A	1.15	0 %	100 %	NA	0 %	0 %	NA
B	1.17	0 %	100 %	NA	0 %	0 %	NA
C	0.34	0 %	100 %	NA	0 %	0 %	NA
D	1.00	0 %	100 %	NA	0 %	0 %	NA
E	0.44	0 %	100 %	NA	0 %	0 %	NA
A-E	0.79	0 %	100 %	NA	0 %	0 %	NA

The outcomes for Organisation III, as approached from PPT-perspectives, are presented in Table 7.9. It is worth noting that the outcomes for these perspectives, namely 0.96, 0.76, and 0.93, respectively, exhibit only small differences.

Table 7.9: Aggregated Outcomes Organisation III: People, Process, Technology

Perspective	*Current Maturity Level*	*Unknown Rate*	*Response Rate*	*Target Maturity Level*	*Unknown Rate*	*Response Rate*	*Maturity Gap*
People	0.96	0 %	100 %	NA	0 %	0 %	NA
Process	0.76	0 %	100 %	NA	0 %	0 %	NA
Technology	0.93	0 %	100 %	NA	0 %	0 %	NA

According to the findings, the current maturity level could be described as approaching the conceptual level. The absence of completing the target maturity states precludes the ability to determine the gaps in maturity.

7.4. Comparative Analysis

The assessment outcomes for the organisations will be compared in this section. Table 7.10 shows an overview of the current maturity levels, target maturity levels, and maturity gaps from the previous section, including colour scales. However, Organisations I and III did not fill in the target maturity states, which consequently made it impossible to determine the maturity gaps. Table 7.10 shows these missing values by *NA* in the cells. Consequently, this implies that only the current maturity states will be compared. To visualise the current maturity levels per dimension, Figure 7.1 and 7.2 are added which respectively present the results in the form of a bar chart and radar chart.

The results demonstrate that **Organisation II** achieved the highest score for all individual dimensions. Furthermore, the second-highest score for all individual dimensions, except for dimension E, was achieved by **Organisation III**. This generally indicates that **Organisation II** is followed by **Organisation III**, with **Organisation I** trailing behind. While **Organisation II** has already reached a maturity level of conceptual, **Organisation III** is approaching this level, while **Organisation I** remains predominantly non-initiated.

Table 7.10: Cross-Case Analysis: Dimensions

Dimension	*Metric*	***Organisation I***	***Organisation II***	***Organisation III***
A	Current Maturity Level	0.54	1.46	1.15
	Target Maturity Level	NA	3.00	NA
	Maturity Gap	NA	1.54	NA
B	Current Maturity Level	0.70	1.63	1.17
	Target Maturity Level	NA	2.94	NA
	Maturity Gap	NA	1.31	NA
C	Current Maturity Level	0.08	1.02	0.34
	Target Maturity Level	NA	2.94	NA
	Maturity Gap	NA	1.92	NA
D	Current Maturity Level	0.67	1.50	1.00
	Target Maturity Level	NA	3.00	NA
	Maturity Gap	NA	1.50	NA
E	Current Maturity Level	0.75	1.46	0.44
	Target Maturity Level	NA	3.00	NA
	Maturity Gap	NA	1.54	NA
A-E	Current Maturity Level	0.48	1.36	0.79
	Target Maturity Level	NA	2.97	NA
	Maturity Gap	NA	1.61	NA

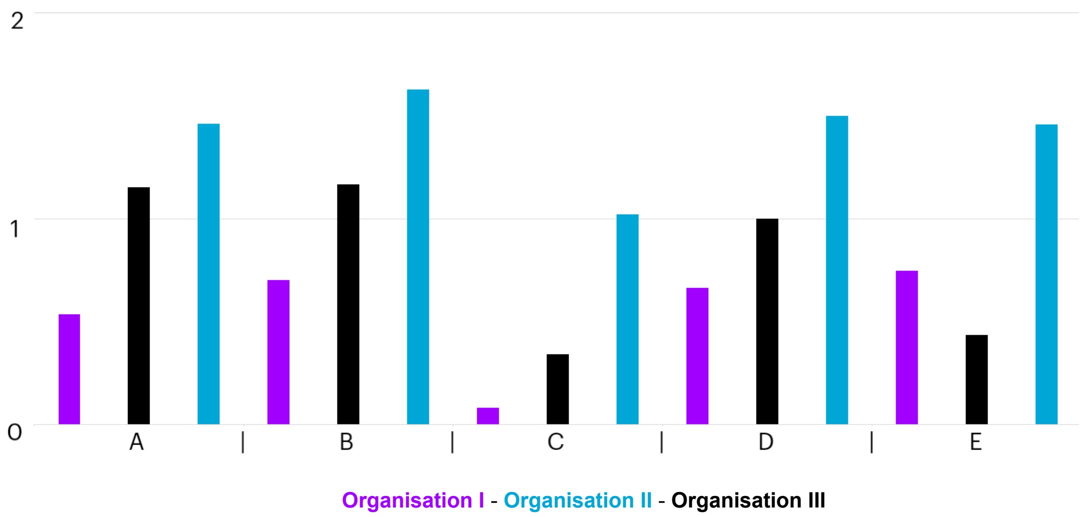


Figure 7.1: Bar Chart: Benchmarking Current Maturity Levels: Dimensions

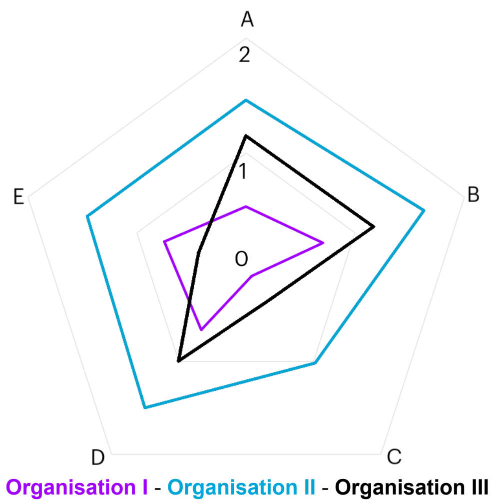
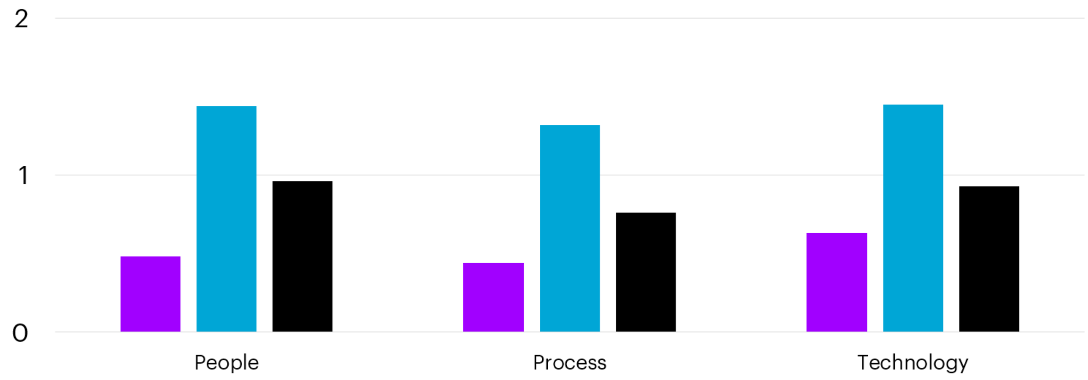


Figure 7.2: Radar Chart: Benchmarking Current Maturity Levels: Dimensions

Table 7.11 shows an overview of the current maturity levels, target maturity levels, and maturity gaps approached from the PPT-perspectives. However, also with regard to the perspectives, only the current maturity levels will be compared. To visualise the current maturity levels per perspective, Figure 7.3 and 7.4 are added which respectively present the results in a bar chart and radar chart. The results demonstrate that **Organisation II** achieved the highest score for all perspectives. Furthermore, the second-highest score for all perspectives was achieved by **Organisation III**. At last, **Organisation I** shows for all perspectives the lowest current maturity levels.

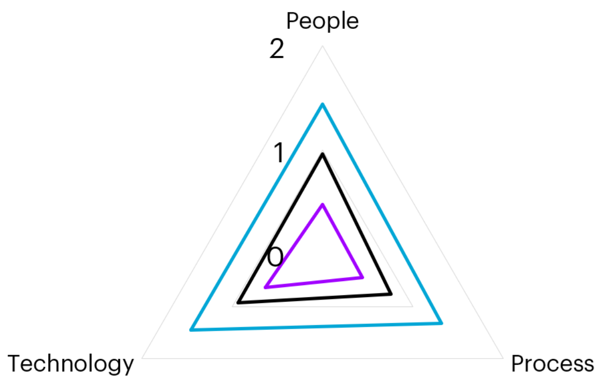
Table 7.11: Cross-Case Analysis: People, Process, Technology

Perspective	*Metric*	***Organisation I***	***Organisation II***	***Organisation III***
People	Current Maturity Level	0.48	1.44	0.96
	Target Maturity Level	NA	3.00	NA
	Maturity Gap	NA	1.56	NA
Process	Current Maturity Level	0.44	1.32	0.76
	Target Maturity Level	NA	2.97	NA
	Maturity Gap	NA	1.65	NA
Technology	Current Maturity Level	0.63	1.45	0.93
	Target Maturity Level	NA	3.00	NA
	Maturity Gap	NA	1.55	NA



Organisation I - **Organisation II** - **Organisation III**

Figure 7.3: Bar Chart: Benchmarking Current Maturity Levels: People, Process, Technology



Organisation I - **Organisation II** - **Organisation III**

Figure 7.4: Radar Chart: Benchmarking Current Maturity Levels: People, Process, Technology

Conclusion Phase D

Phase D aimed to observe the applicability and usefulness of the DMMAM. The term applicability pertained to whether it is practically feasible to demonstrate the DMMAM, while usefulness referred to the ability to provide the desired outcomes regarding the proposed maturity metrics and charts. This section will provide the conclusion on whether this objective is achieved by demonstrating the DMMAM in various cases. A case was defined as the process of conducting a data mesh maturity assessment for an organisation.

To conduct maturity assessments in cases, the involvement of organisations was necessary. Three organisations were identified and engaged through personal outreach facilitated by Accenture's professional network. These organisations were all large multinational businesses that hold leading positions in their respective industries. The selection of these organisations was based on their exploration or implementation of data mesh. In addition, the objective was to include three participants from each organisation who collectively possessed the knowledge to answer all the questions in the assessment. Whereas Organisations I and II were represented by three participants each, Organisation III was represented by only two participants. All eight participants held roles within the field of data management and possessed 10 to 20 years of working experience. Within each organisation, one participant was assigned as the organisational representative who acted as the spokesperson during the execution of the case.

The case followed a structure consisting of three assessment activities, as outlined by DAMA International (2017). These activities encompassed planning the assessment activities, conducting the maturity assessments, and interpreting the results. Firstly, the case commenced with the planning, which involved an introductory meeting with the organisational representative. During this meeting, the model and assessment were explained, and the questionnaire was shared. The organisational representative then internally distributed the questionnaire to the participants. Secondly, the participants individually completed the self-assessments. Throughout the assessment process, participants had the option to schedule a discussion session with the designer if any uncertainties arose. However, it turned out that such sessions were not required for any of the cases. Lastly, the completed assessment forms were collected and analysed by the designer. The analysis focused on calculating the maturity scores for the specified metrics and creating charts. The results were subsequently shared with the organisational representative and presented in a closing workshop for each organisation which all participants attended. This workshop provided an opportunity to discuss and delve into the results. The three cases were conducted in an identical manner.

The outcomes of the maturity assessment revealed that all organisations are still in the early stages of their data mesh implementation. The current overall data mesh maturity scores for Organisation I, II, and III were 0.48: *Non-Initiated*, 1.36: *Conceptual*, and 0.79: *Conceptual*, respectively. Moreover, the assessment showed that Organisations I and III did not provide information on their target maturity states due to organisation-specific reasons. As a result, it was not possible to determine their specific maturity gaps. In contrast, Organisation II set a target overall data mesh maturity score of 2.97, which reflects a maturity gap of 1.61. This indicates that Organisation II is striving for the highest level of data mesh maturity, which is *Achieved*. The existence of this maturity gap suggests that there is still a large distance to cover to reach the desired level.

After the assessment, a comparison of outcomes was conducted across the organisations. The analysis indicated that Organisation II has surpassed the maturity level classified as *Conceptual*, followed by Organisation III which is approaching this level. In contrast to Organisation I, which remains predominantly *Non-Initiated*. The analysis further demonstrated that this order of maturity levels was also reflected across the individual dimensional scores and the scores from PPT-perspectives.

In conclusion, after conducting the cases for demonstration, the designed DMMAM showed that it was both applicable and useful. The demonstration confirms its practical feasibility for conducting a data mesh maturity assessment, establishing its applicability. Furthermore, the model proved to be useful as it provided the organisations with the outcomes for the proposed maturity metrics and charts by following the selected assessment activities. However, it should be noted that the remark arises from the fact that Organisations I and III did not fill in target maturity states. Nonetheless, as Organisation II was able to provide this information, and Organisations I and III indicated that this issue was unrelated to the designed DMMAM itself, it will be concluded that the DMMAM is also perceived as useful.

Phase E

Evaluation

8

Model Evaluation

While Phase D concluded that the designed DMMAM was applicable and useful, Phase E will continue by evaluating the extent to which this was the case. This also implies that feedback from the assessment participants on how the cases have progressed will be evaluated in this phase. To do this, the reliability and validity of the model and the assessment activities will be examined. The deliverable in this phase will be the final DMMAM. The sub-question, sub-objective, and sub-deliverable of this phase are provided below.

Sub-Question 4

To what extent is the designed data mesh maturity assessment model applicable and useful?

Sub-Objective 4

Concluding about the extent of applicability and usefulness of the data mesh maturity assessment model.

Sub-Deliverable 4

Final data mesh maturity assessment model.

Chapter 8 evaluates the DMMAM in terms of reliability in Section 8.1 and validity in Section 8.2.

8.1. Reliability

Reliability refers to whether the assessment model results are consistent over multiple assessments under the same conditions (Creswell, 2018). The reliability will be evaluated by looking into criteria as defined by Gökalp et al. (2022). Gökalp et al. used criteria to evaluate their designed big data analytics process capability assessment model. These evaluation criteria are also expected to be relevant in this research since these are defined for the purpose of reflecting on a data maturity assessment model. In this research, the criteria are divided into two categories: those related to the applicability and usefulness of the model, and those related to the applicability and usefulness of conducting the assessment. Table 8.1 provides an overview of the criteria with descriptions. The criteria will be evaluated with respect to their impact on the reliability in Section 8.1.1 and 8.1.2.

Table 8.1: Reliability Criteria

Nr.	*Criteria*	*Description*
Model		
1	Fitness for Purpose	The level of fitness to guide an organisation during the implementation process.
2	Completeness	The level of completeness of dimensions and characteristics in addressing all aspects of data mesh.
3	Granularity of the Dimensions	The level of detail in terms of the characteristics of each dimension.
4	Definition of the Measurement Attributes	The level of detail in terms of the maturity levels.
5	Description of the Assessment Model	The level of detail regarding the assessment method.
6	Objectivity of the Assessment Model	The level of objectivity of the assessment method.
Assessment		
7	Audience	Organisations and participants fitness for purpose.
8	Verification	Case structure in terms of the self-assessment questionnaires, individual discussion sessions, and closing workshops.

8.1.1. Model Reliability

Fitness for Purpose

This criterion refers to the extent to which the designed DMMAM effectively guides an organisation during the implementation of data mesh. Various reasons, as presented in Phase B, showed that the designed model fits to meet this purpose. First of all, the maturity assessment model shows, by presenting the dimensions and characteristics, what data mesh entails. This overview provides an understanding of which aspects need to be taken into account during the implementation. Secondly, the criteria and requirements for all characteristics over the levels present the path to a mature implementation. Thirdly, by asking for assessing the current and target maturity states, the path to opportunities is determined. Altogether, the assessment model explains what a mature data mesh means and provides the phases from an immature to a mature implementation. Therefore, the designed DMMAM is considered fit for its purpose to provide the needed guidance during the data mesh implementation process.

The model is designed to only meet this purpose of providing guidance to implementing data mesh. This means that when the model will be used for other purposes, reliability is not guaranteed. To illustrate, Al-Sai et al. (2022) mentioned that the maturity assessment model would help in assessing the readiness for data mesh by assessing current maturity states. However, an maturity assessment model is not the same as a readiness assessment model. According to DAMA International (2017), a maturity assessment model evaluates an organisation's maturity levels, whereas a readiness assessment model evaluates an organisation's readiness as preparation to anticipate an upcoming project or initiative (Barham & Daim, 2020). The emphasis in a readiness assessment lies in the preparation of an initiative, rather than on its development once it has been implemented. Using the DMMAM solely as a readiness assessment would not align with the intended purpose of the DMMAM-design in this research. As an additional example, it is not recommended to use the DMMAM to evaluate data management maturity, as the DMMAM does not encompass the relevant data management aspects. In conclusion, the consistency of the results obtained from the DMMAM across multiple assessments is consistent with maintaining adherence to the model's intended purpose.

Completeness

The level of completeness refers to whether the dimensions and characteristics address all the aspects of the data mesh. In this research, the DMMAM is considered as complete as soon as it covers the four data mesh dimensions introduced by Dehghani (2022a). In addition, the characteristics need to cover all relevant aspects according to the four dimensions. There is no fixed number of characteristics to include to obtain completeness. The number of characteristics is dependent on the research scope and the level of detail required to have a successful assessment.

The designed DMMAM in this research evaluates data mesh by five dimensions, including the four as introduced by Dehghani. The fifth dimension was added to ensure all essential data mesh characteristics were able to be classified into the dimensions. Due to the designed DMMAM includes the four main principles, the DMMAM is considered as complete. Additionally, the established 54 characteristics are expected to be sufficiently collectively exhaustive in assessing data mesh. Nevertheless, the level of completeness could be improved by adding more and more characteristics. However, increasing more characteristics would at the same time impact the assessment feasibility and increases the possibility of having overlap among the characteristics. Overall, 54 characteristics are perceived as a well-considered and extensive set.

In terms of reliability, the consistency of assessment outcomes across multiple assessments depends on whether all characteristics were evaluated during each assessment. This does not imply that all participants are required to evaluate all characteristics, but rather that the participants from one organisation collectively were able of responding to all questions. If re-assessments are conducted at fixed intervals, the results could only be compared reliably when all characteristics were evaluated during each assessment to ensure that all relevant aspects were covered.

Granularity of the Dimensions

In terms of granularity, the DMMAM is considered as sufficiently detailed to cover all relevant aspects of data mesh. The level of detail is reflected by the number of characteristics per dimension and by the definitions.

Dimensions A, B, C, D, and E consist of respectively 13, 9, 16, 8, and 8 characteristics. It was needed for some dimensions to include more characteristics to ensure all relevant aspects were covered. To illustrate, dimensions D and E contain 8 characteristics, whereas dimension C is reflected by 16 characteristics. The relatively large number of characteristics for dimension C is mainly because the 7 usability attributes are included individually instead of collectively. However, this does not mean that the number of characteristics also represents the importance of a dimension. The reason why not all dimensions have an equal number of characteristics is that some dimensions are represented by more characteristics based on the availability of research that has been conducted.

Next to the number of characteristics, the definitions also impact the level of detail. To illustrate, characteristic B4: *Ownership* refers to whether a *domain owner is defined and allocated*, whereas characteristic E5: *Standardisation* is defined as *enforced global standards to enable interoperability, such as consistent processes and tools; global definitions for domains, products, and ownership in the glossary; data product business and technical KPI's; entity ID's (polysemes identifiers); data sharing API's; common metadata fields including SLO's, documentation, and modelling language; schema linking; data linking; schema stability (backward compatibility)*. These definitions show the difference in the level of detail. Characteristic E5 could have been split into more characteristics. However, the design process involved iterating to strike a balance between creating a valuable model, while keeping it feasible. In the situation of characteristic E5, providing various examples showed more detail compared to the definition of characteristic B4. The inequality in the level of detail across different characteristics was inevitable and is not considered as problematic. The intention was that even the least described characteristics were still sufficiently detailed to express and convey their meaning to the participants.

The inability to accurately express and convey the meaning of data mesh dimensions through characteristics and definitions would increase the possibility of obtaining unreliable results. Inadequate detail could result in varying interpretations among participants, potentially impacting the consistency of results across multiple assessments.

Definition of the Measurement Attributes

Defining the measurement attributes refers to the level of detail in terms of the maturity levels. Four maturity levels were established to evaluate the data mesh maturity for an organisation. The findings from the benchmark in Section 3.3 showed that 4, 5, or 6 levels are common in data maturity assessment models. In this research, four maturity levels have been selected since it is expected by the designer that from four maturity levels, the data mesh implementation process could be provided in a way it would be helpful as guidance for organisations. More specifically, the objectives analysis showed that the top-level objective is obtaining a successful DMMAM, meaning that it aims to achieve the goals as has been concluded in Chapter 3. It is expected that a four-level DMMAM contributes to discovering and evaluating the characteristics of data mesh, educating the users about the concepts, roles and responsibilities, and it would enable the establishment of roadmaps after the assessment. Moreover, by adding more levels, it would become harder to distinguish the different levels for the designer, user, and client. Adding more levels necessitates greater nuance in the criteria and requirements used to differentiate between the various levels. Due to data mesh being a relatively new concept, the designer, users, and client will not always be able to state and note the differences across the definitions while having more maturity levels. Thus, adding more maturity levels would adversely impact reliability. When the designer is not able to differentiate the maturity levels, overlap will arise between the descriptions. In case the participant is not able to distinguish the maturity levels, the possibility of incorrectly evaluating the maturity score will increase. In this situation, performing multiple assessments will not result in consistency in the maturity outcomes. For this reason, four maturity levels provide a sufficient and reliable starting point for reflecting on the data mesh implementation process.

To increase the reliability, each numerical maturity level has a label and definition provided to express and convey their meaning to the participants. The definitions are based on three factors: the degree to which the data mesh characteristics are implemented, the level of understanding of the data mesh principles, and the level of experience with the data mesh implementation. These three factors are expected to encompass the various data mesh elements of maturity levels. The combination of the well-considered number of maturity levels and the detailed descriptions contribute to the reliability of the model.

Description of the Assessment Model

The objectives analyses in Section 5.1 showed that the model needs to be *unambiguous*, *explainable*, *self-describing*, *well-defined*, *well-articulated*, *recognisable*, and *tailored*. All these objectives could impact the model's reliability. These objectives have been taken into account regarding the description of the assessment method in several ways. First of all, the following introduction was provided to inform the participant about the assessment model purpose and the categorisation of the assessment into dimensions with characteristics:

"This document provides a data mesh maturity assessment model. The model is designed to enable the assessment of how mature a data mesh implementation is within an organisation. Five dimensions are considered to assess the data mesh maturity. These five dimensions are represented by, in total, 54 characteristics."

Secondly, the model presents the following instructions for the self-assessment:

"Please rate the following 54 questions on the current and target level of maturity that you find most fitting for your organisation. Each statement should be rated on a scale from Level 0. "Non-Initiated" to Level 3. "Achieved", or "Unknown". The descriptions guide you in making your choice. If you have any comments, please enter them in the comments section."

Thirdly, control questions are included to guide the participant conveniently through all 54 characteristics. Fourthly, at least one well-defined criteria or requirement is included for the different maturity levels. Fifthly, for all 5 dimensions and 54 characteristics are definitions included in the document.

Overall, incorporating an introduction, instructions, control questions, criteria and requirements, and definitions contribute to the set of objectives and reduces the bias between the designer's intended meaning and the participant's interpretation. As a result, it is expected that results over multiple assessments will therefore be consistent.

Objectivity of the Assessment Model

The assessment model aims for an objective measurement of the data mesh maturity. Including objective measures would improve the consistency of results over multiple assessments. Objectivity has been taken into account in several ways. First of all, the dimensions, characteristics, and maturity levels which need to be evaluated are as much as possible objectively described and specified. Secondly, each dimension is represented by 8 to 16 characteristics. Having too few characteristics would result in the possibility that one single characteristic will have a large impact on the final dimensional maturity score. By incorporating a larger number of characteristics, the dimensions would become less susceptible to single extreme values. Thirdly, the assessment structure is standardised to avoid variety in outcomes due to different assessment conditions. Fourthly, the constraints analysis presented that at least three participants are required to fill in the questionnaire. Having too few participants would increase the overall impact of individual bias. In addition, it is required to have a balanced group, such that the participants together are able to answer all questions. Fifthly, the questions are formulated objectively as much as possible. Furthermore, the criteria and requirements are also formulated in a measurable manner as far as possible. Sixthly, the participants had the possibility to schedule a discussion session with the designer in case they would have had questions or remarks about the assessment. In addition, the designer's email address was provided in case of brief questions or remarks. Seventhly, a small survey was added to the assessment where the participant filled in the extent to which everything was *well-explained*, *well-defined*, and *self-describing* according to their opinion. The answers would help the designer to assess whether additional instructions are needed. Lastly, the maturity scores were calculated by a fixed formula where all characteristics had equal weights. Assigning different weights by default would have been arbitrary. Altogether, these design choices reflect the aim to have an objective measurement. Keeping these conditions consistent over multiple assessments would guarantee reliability in the outcomes.

However, subjectivity still appears in the assessment. To illustrate, characteristic C9 refers to the *Lead Time* to successfully build, test, deploy, discover, use, or change data products or policies. Maturity levels 0, 1, 2, and 3 provide the following descriptions:

Level 0	Level 1	Level 2	Level 3
Relatively high overall lead time for building, testing, deploying, discovering, using, or changing data products or policies. All processes are still manually and there is a high dependency on the central IT or Analytics team to perform tasks, and get access or approvals.	High overall lead time for building, testing, deploying, discovering, using, or changing data products or policies. However, processes are getting more automated. Central IT or Analytics team is still performing most tasks and providing access or approvals.	Reduced overall lead time to successfully build, test, deploy, discover, use, or change data products or policies. Increased level of domain autonomy, due to improved automated processes and training. Domains operate more autonomously and are less dependent on central IT or Analytics team.	Relatively low overall lead time to successfully build, test, deploy, discover, use, or change data products or policies. Processes are completely automated. Domains operate autonomously and are not or only less dependent on central IT or Analytics team.

These established descriptions are prone to subjectivity. To increase objectivity, this characteristic could have been broken down into six distinct characteristics: the lead time for *building* data products, the lead time for *testing* data products, the lead time for *deploying* data products, the lead time for *discovering* data products, the lead time for *using* data products, and the lead time for *changing* data products. Furthermore, the descriptions could have been objectified by including quantitative measures. To illustrate, the description for discovering a data product according to maturity level 0 would become:

"Discovering a data product or its policies would take more than one day. It is not possible to search for data products by the domains themselves. A central IT or Analytics team needs to be consulted to ask which data products exist."

Despite this adjustment would objectify the assessment, it is not desirable to have this degree of objectivity for several reasons. First of all, splitting the lead time into six individual measures would heavily increase the number of characteristics. Splitting all characteristics into orthogonal measures would easily double the total amount. As a result, the assessment would become too extensive and time-consuming. Secondly, the insight that the lead time is an important characteristic to take into account to measure the data mesh maturity is in general more important than the exact time it takes to discover a data product. Due to data mesh is still in its infancy, it is expected that organisations will not be able to answer the questions in this level of detail. Thirdly, having measurable descriptions does not always mean that it would be a more reliable measure. To illustrate, characteristic A2: *Culture, Mindset,*

& Values refers to *the self-serve culture in which people understand the importance of data, are data literate, have a product-sharing mindset, trust the validity of data, and accomplish data use cases*. The question of whether an organisation accomplished a cultural shift from data protection towards data-sharing is subjective in nature. In this situation, it is undesirable to quantify characteristic A2: *Culture, Mindset, & Values* to obtain objectivity. Fourthly, several data mesh concepts are still abstract or complex. In such cases, having subjective measures would be more appropriate. To illustrate, characteristic D1: *Infrastructure & Platform* refers to *the established infrastructure as a platform which provides the set of technologies to enable domain teams to search and share data products, to create and set policies for data products, and for storing, computing, and caching purposes*. What kind of technologies are considered, whether cloud usage is advantageous, and how these would be integrated have not yet been researched.

To conclude, objective measures have always been the starting point. However, the model should not be overwrought with unnecessary details or complexity which would adversely impact the assessment understandability and client-friendliness. In addition, intangible and abstract characteristics should not be objectified. Eventually, a balanced set of objective and subjective measures has been established.

8.1.2. Assessment Reliability

Audience

The audience for the designed DMMAM refers to organisations which are considering or have started implementing data mesh. For those organisations would a data mesh maturity assessment be helpful as guidance to discover, start, or further plan the data mesh implementation. According to Participants B, D, I, and N, as presented in Table B.18, data mesh does not vary in its implementation across industries. The designed model could therefore be used for assessing organisations from different industries.

Participants B, C, D, E, F, G, and N, as presented in Table B.7, argued that the organisational participants must be balanced as a group to come to reliable results. It is needed to have both data providers and data consumers involved to cover the complete data supply chain. Moreover, it is needed to ensure sufficient expertise is available to address all the details. Some examples of roles within an organisation that come to mind include data scientists, IT-architects, data stewards, data custodians, data privacy officers, information security specialists, and data governance specialists. In addition, data leaders, managers, as well as executives need to be involved to include all hierarchical perspectives. Whereas data practitioners will evaluate data mesh from a more operational perspective, managers and executives would approach it more strategically. Moreover, including the hierarchical perspectives also contribute to having a group which varies in terms of years of experience. At last, next to involving technical stakeholders, business stakeholders also needed to be included. Examples of roles are business analysts and product owners.

In terms of reliability, this composition of participants is needed for correctly assessing the data mesh maturity. As a result, it is expected that twenty to thirty participants are needed in each assessment to cover all perspectives. This has not been met in the cases, which also means not all aspects have been taken into account to a sufficient extent. While conducting re-assessments, the outcomes will only be consistent as long as the group is balanced each time.

Verification

The applicability and usefulness of the designed DMMAM are evaluated in cases for three organisations. To enable a reliable comparison of the results, the assessment structure, as presented in Section 7.2, was identical in all cases. However, the only remark to be made here is that Organisation III was represented by two participants, whereas Organisations I and II had three participants involved each.

Overall, evaluating the criteria by Gökalp et al. highlighted the design considerations which impacted the reliability of the model and assessment. In conclusion, the designed DMMAM is considered as reliable to guide organisations during the data mesh implementation process. However, it is important to take into account that this model is the first attempt to provide a comprehensive framework to assess the data mesh maturity for organisations. As such, the assessment model is applicable and useful for all organisations which are considering or started the implementation of data mesh and would need initial guidance in terms of a maturity assessment model. Nevertheless, it is important to acknowledge that the model is not without limitations. The concept of data mesh is still in its infancy, which means that the model would need ongoing refinement and improvement.

8.2. Validity

According to Creswell, validity reflects on whether the results represent what they are supposed to measure. In other words, it refers to whether the model is accurate in assessing the data mesh maturity. According to Peffers et al. (2007), evaluating the model entails comparing the objectives to the observed outcomes from the demonstrations in Phase D. In addition, Peffers et al. indicate that a satisfaction survey or client feedback is helpful in obtaining an understanding of the extent of applicability and usefulness. This section will therefore evaluate in Section 8.2.1 five objectives by asking the participants from the cases to provide feedback. Evaluating the constraints and functional analyses in respectively Section 8.2.2 and 8.2.3 will also help in assessing the impact of individual design choices on the model validity.

Table 8.2 presents the aggregated responses from all organisations on the feedback questionnaire. In total, this questionnaire is filled in by eight participants. The first question asked how much time it took the participants to fill in the assessment. The second question consists of six statements which the participants assessed to what extent they agreed or disagreed with the statements. The results show that filling in the self-assessment took the participants an average of 50 minutes. Based on the individual results, it was found that the assessment lasted between 30 and 90 minutes for the participants. Out of the eight participants, only one exceeded the time constraint of 60 minutes. The second question shows the number of participants who provided a specific answer for each statement, with the total for each row adding up to eight. The last statement pertains to whether the assessment model is helpful for the organisation. Except for one participant who chose *Disagree* and one participant who chose *Neutral*, it was found that six out of the eight participants chose *Agree*. Therefore, the assessment was generally perceived as helpful by the participants. The findings from the other five statements will be explained in Section 8.2.1.

Table 8.2: Aggregated Feedback Questionnaire

1. How much time did it take you to fill in this assessment?

50	minutes

2. To what extent do you agree with the following statements?

Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree

	Statement	SA	A	N	D	SD
1	All aspects of data mesh are covered.		4	4		
2	All maturity levels were clearly distinguished.		5	3		
3	Everything was explained well.		7		1	
4	Everything was well-defined.		4	1	3	
5	The assessment was self-describing.	1	5	1	1	
6	The assessment is helpful for us.		6	1	1	

8.2.1. Objectives Evaluation

The objectives analysis in Section 5.1 presented sixteen bottom-level objectives, as presented in Table 8.3. To evaluate whether the model achieved its objectives during the cases, five objectives are selected to be evaluated in terms of validity. It was decided not to evaluate all sixteen objectives with the participants, as they also have limited time available. In addition, these specific five objectives were chosen because it was expected that participants would have a good understanding of what these objectives meant.

Table 8.3: Objectives Evaluation

Nr.	*Objective*	*Description*
1	Pragmatic	Solving problems in a sensible way that suits the conditions that really exist now.
2	Actionable	Able to be used as a reason for doing something. Outcomes need to be translated into actionable next steps.
3	Marketable	Able or fit to be sold or marketed.
4	Durable	The model should be able to continue to exist for a long time, by being maintainable and sustainable.
5	Complete	Data mesh needs to be approached from all the different perspectives.
6	Orthogonal	Independent, no overlap in dimensions, characteristics, and maturity levels.
7	Unambiguous	Expressed in a way that makes it completely clear what something means.
8	Measurable	Aspects need to be measurable to have a correct assessment.
9	Explainable	Model and assessment outcome should be understood by the user and client.
10	Consistent	Always behaving in a similar way. The outcome should be consistent regardless of who the user or client is.
11	Achievable	An assessment needs to be possible in time and resources.
12	Self-Describing	Serving to describe oneself. The user should be able to answer the questions without any help from the client.
13	Well-Defined	Clearly expressed, explained, and described dimensions, characteristics, and maturity levels.
14	Well-Articulated	Able to express meanings easily and clearly and show their quality.
15	Recognisable	Concepts needs to be familiar to user and client.
16	Tailored	To adjust or expand something to the specific needs of the user and client.

Complete – The first statement in the survey asked whether all aspects of data mesh are covered in the assessment. The results show that four participants answered *Neutral* and four participants filled in *Agree*. Taken together, these findings collectively demonstrate a slightly positive outcome. It is expected that four participants have chosen *Neutral* as they are unsure whether all aspects are adequately covered, due to data mesh is still relatively unfamiliar to many.

Orthogonal – The second statement asked whether all maturity levels were clearly distinguished. Three participants filled in *Neutral* and five participants filled in *Agree*. Also, this statement shows a collectively slightly positive outcome.

Explainable – The third statement asked whether everything was explained well. Except for one participant who chose *Disagree*, seven participants selected *Agree*. This result indicates that almost everyone agrees that everything was *explainable*. Among all statements, this question has received the most uniform responses.

Well-Defined – The fourth statement asked whether the assessment was well-defined. Three participants filled in *Disagree*, one participant is *Neutral*, and four participants filled in *Agree*. These results demonstrate the greatest division, with a few individuals disagreeing, but ultimately the majority agreeing. Since data mesh is still relatively novel and has been scarcely researched, the DMMAM intentionally includes several descriptions that are abstractly formulated. As a result, it is expected that therefore a large number of participants have chosen *Disagree* for this statement.

Self-Describing – The fifth statement asked whether the assessment was self-describing. Four answers were selected by the participants. For the answers *Disagree*, *Neutral*, and *Strongly Agree* were one participant who selected this option. In addition, five participants chose *Agree*. This indicates an overall positive view on this matter.

To conclude, the five statements were evaluated to gain insight into whether the DMMAM has achieved its intended objectives, as assessed by the model's users. It could be concluded that the results are generally positive. In total, *Agree* and *Strongly Agree* were selected 26 times. *Disagree* was chosen only 5 times, and *Strongly Disagree* was not selected at all. These results reflect the fact that the model objectives are generally validly achieved.

8.2.2. Constraints Evaluation

The constraints analysis in Section 5.2 presented six limits, as provided in Table 8.4, which the design must meet. The first constraint is selected to be discussed in terms of validity. The remaining constraints have already been thoroughly discussed in Section 8.1, except for the sixth statement regarding the discussion session. However, this is not relevant to discuss in this context since no discussion session was requested by the users.

Table 8.4: Constraints Evaluation

Nr.	*Constraint*
1	Data mesh characteristics must be mutually exclusive and collectively exhaustive.
2	Number of maturity levels.
3	Number of characteristics and questions must balance completeness and research capacity.
4	Number of user participants.
5	Represented user participants must be balanced as group.
6	Duration user participant discussion session.

Collectively exhaustiveness refers to the objective *completeness*, i.e., that all data mesh aspects must be covered. Mutual exclusiveness refers to the objective *orthogonality*, i.e., ensuring that there is no overlap or correlation amongst the aspects. Since *completeness* is discussed in Section 8.1.1, this section will focus on orthogonality. The implication of not meeting orthogonality implies that the characteristics would not all uniquely contribute to the overall score (Little et al., 2006). In other words, some effects will be double-counted. Therefore, it is important to meet this constraint to ensure optimal model validity. This section will evaluate to what extent the set of five dimensions, represented by 54 characteristics, overlap each other and what actions were taken to reduce this overlap.

Figure 8.1 shows by arrows that the initial four dimensions are all interrelated with each other (Dehghani, 2022c). More specifically, the direction of the arrow represents that the implementation of the *from* dimension creates the challenge that the *to* dimension addresses. This interrelation includes that the characteristics belonging to the dimensions are also dependent on each other. This violates the constraint of mutual exclusiveness. Furthermore, in this research, an extra dimension is added to cover the data foundation and organisational change aspects. Figure 5.3 showed dimension A as the baseline to implement dimensions B, C, D, and E. This also means that there is a dependency between the characteristics from dimension A in relation to the others. In short, the overlap is present and unavoidable due to data mesh being introduced as the interrelation of dimensions. As a result, the constraint set in this research needs to be relaxed.

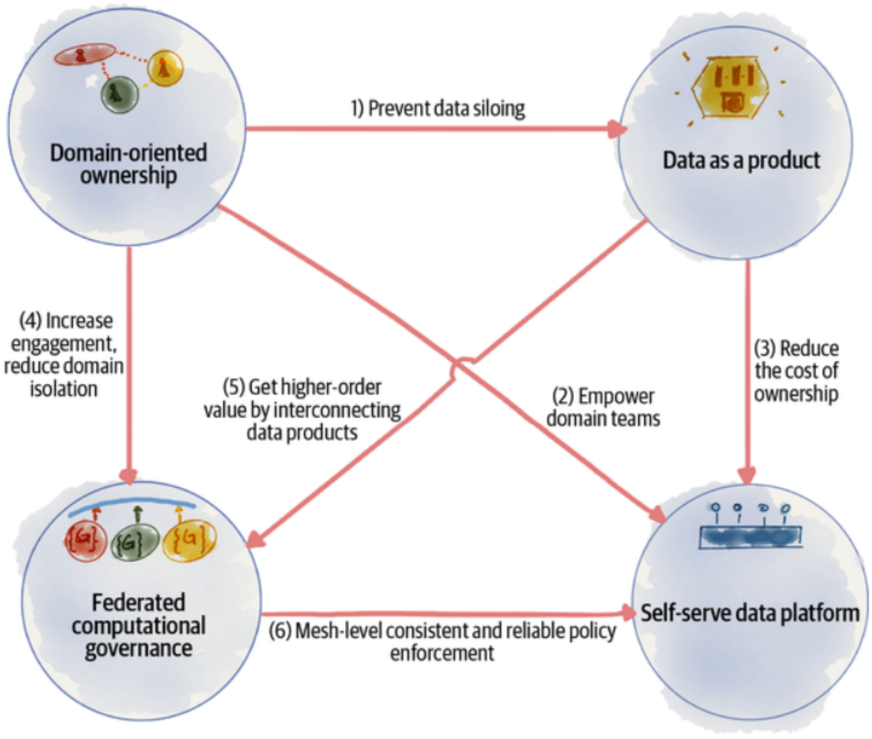


Figure 8.1: Interplay Dimensions

To illustrate where the overlap is still present in the characteristics from the DMMAM, Table 8.5 shows numerous examples, which will be explained.

The overlap between A2: *Culture, Mindset, & Values* and A4: *Curiosity & Ability* is centred around the importance of creating a data-driven culture in which individuals and teams are encouraged to explore and experiment with data.

Table 8.5: Correlation Characteristics

ID	*Characteristic*	*ID*	*Characteristic*
A2	Culture, Mindset, & Values	A4	Curiosity & Ability
A3	Value Realisation	A9	Value Adding Use Cases
A8	Change Management	A13	Training
A10	Roles	B4 C2 D4 D5 E7	Ownership Ownership Ownership Platform Team Governance Team
B5	Autonomy	A11 B6	Skills & Capabilities Cross-Functional Teams
C4	Production & Sharing	B8 B9	Producers Consumers
C15	Security	E1	Security & Compliance
C3	Discovery Tool	C16	Accessibility
C9	Lead Time	D3	Performance
E2	Global Policies	E5	Standardisation
E5	Standardisation	B1 C1 C13 C14	Definition Definition Descriptiveness Interoperability

The overlap between A3: *Value Realisation* and A9: *Value Adding Use Cases* is centred around identifying and delivering value out of data. *Value Realisation* emphasises the need to track distinct value created by data & analytics and machine learning. Similarly, *Value Adding Use Cases* highlights the importance of identifying use cases for data products that create inherent value. Despite *Value Adding Use Cases* focusing more specifically on offering cross-domain value by data products, both definitions focus on value realisation. Additionally, both *Value Realisation* and *Value Adding Use Cases* suggest that tracking the value is important for demonstrating their impact.

The definitions of A8: *Change Management* and A13: *Training* focus on education and awareness-raising, as well as the development of new skills and capabilities across the organisation. However, *Change Management* addresses change more holistically, whereas *Training* focuses solely on educational programmes.

The overlap between A10: *Roles*, B4: Domain *Ownership*, C2: Data Product *Ownership*, D4: Platform *Ownership*, D5: *Platform Team*, and E7: *Governance Team* is centred around defining and allocating roles and responsibilities. *Roles* emphasises the need to define specific roles and teams in general. Domain *Ownership*, Data Product *Ownership*, Platform *Ownership*, *Platform Team*, and *Governance Team* are defined in more detail to obtain an understanding of to what extent the roles are defined.

The similarity between B5: *Autonomy*, A11: *Skills & Capabilities*, and B6: *Cross-Functional Teams* refers to establishing autonomy by cross-functional teams which require business, technology, and data skills and capabilities.

The overlap between C4: *Production & Sharing*, B8: *Producers*, and B9: *Consumers* addresses the importance of consuming, producing, and cross-domain data product sharing to create a more collaborative ecosystem.

The overlap between definitions C15: *Security* and E1: *Security & Compliance* is in the aspect of security policies.

Both definitions address the importance of having policies in place to ensure that data products are secure and comply with regulations. However, *Security* refers to the policies in the data products, whereas *Security & Compliance* refers to the global governance policies.

There is an overlap between C3: *Discovery Tool* and C16: *Accessibility* as they both focus on enabling access to data products. *Discovery Tool* includes the establishment of a central data product discovery tool, which enables a seamless way to search, explore, request, and share relevant data products. Similarly, *Accessibility* presents the importance of enabling the possibility that various users are able to access and shop data products. Both of these definitions emphasise the importance of making data products easily accessible to users.

C9: *Lead Time* and D3: *Performance* both focus on measuring the effectiveness and efficiency of the data platform. Whereas *Lead Time* looks at the lead time for building, testing, deploying, discovering, using, or changing data products or its policies, *Performance* takes into account various factors such as platform costs, user experience, number of people using the platform, level of automation, and the number of features available.

The overlap between the definitions for E2: *Global Policies* and E5: *Standardisation* refers to the establishment of global standards to enable data product interoperability. *Global Policies* focuses on global communication protocols and standards that cover the full scope of central governance. Similarly, *Standardisation* is defined as the enforced global standards to enable data product interoperability. Both definitions highlight the need for establishing and enforcing global standards and protocols to enable seamless communication and data product exchange.

The similarity amongst E5: *Standardisation* and B1: Domain *Definition*, C1: Data Product *Definition*, C13: *Descriptiveness*, and C14: *Interoperability* is centred around the need for standardisation to improve the understandability by documentation about what data domains and data products entail.

In conclusion, this section emphasises that overlap is still available, but at the same is assumed as unavoidable. Additional revisions could attempt to reduce the overlap and thereby increase the DMMAM validity. In general, the DMMAM is still deemed to be sufficiently valid for its purpose. The definitions of all the characteristics are formulated with the aim of delineating the aspects, thereby ensuring that the participants comprehend the distinctions. This would serve to mitigate the probability that a participant may conflate multiple characteristics as a single measure.

8.2.3. Functions Evaluation

The functional analysis in Section 5.3 presented four functions, as provided in Table 8.6, the design must *do*. The function referring to providing data mesh maturity scores for the different data mesh dimensions will be evaluated in terms of validity.

Table 8.6: Functions Evaluation

Nr.	*Function*
1	Providing overall data mesh maturity score.
2	Providing maturity scores for the different data mesh dimensions.
3	Providing data mesh maturity scores from People, Process, Technology perspective.
4	Providing guidance for achieving higher levels of maturity.

Given the interrelations among the different dimensions, it raises the question of whether it would be valid to score the dimensions individually and whether it would be valid to conduct a smaller assessment by only examining the characteristics of a single dimension. Providing individual maturity scores would still be expected to be beneficial for the participating organisation to see which dimensions are more mature than others. In addition, performing re-assessment at regular intervals would provide insights into the progress organisations make with respect to the dimensions individually. However, the overall data mesh score would be relatively more valid than the individual scores, due to the overlap among the characteristics across the dimensions. Secondly, it is discouraged to selectively assess a subset of characteristics, given data mesh is introduced as the composition of interdependent dimensions.

Conclusion Phase E

Phase E aimed to conclude the extent of applicability and usefulness of the DMMAM. This involved an examination of both reliability and validity aspects.

Reliability, as defined by Creswell (2018), refers to the consistency of assessment results across multiple assessments conducted under the same conditions. The reliability of the model and the assessment were assessed by six and two criteria respectively, which were introduced by Gökalp et al. (2022) to evaluate the applicability and usefulness of their data maturity assessment model.

The model reliability was evaluated by looking into its *Fitness for Purpose*, *Completeness*, *Granularity of the Dimensions*, *Definition of the Measurement Attributes*, *Description of the Assessment Model*, and *Objectivity of the Assessment Model*. First of all, *Fitness for Purpose* refers to the extent to which the designed DMMAM effectively guides an organisation during the implementation of data mesh. It involves demonstrating what data mesh entails by presenting dimensions and characteristics, outlining the criteria and requirements for a mature implementation, and identifying areas for improvement. It has been concluded that the reliability of the assessment could only be assured when the DMMAM is used for this purpose. Secondly, *Completeness* refers to whether the dimensions and characteristics of the DMMAM adequately cover all aspects of data mesh. The DMMAM is considered complete because it incorporates the four main dimensions introduced by Dehghani (2022a). Additionally, the established 54 characteristics are expected to collectively cover all relevant aspects. The reliability of the assessment could be assured as long as all the characteristics are adequately covered during each assessment. Thirdly, regarding the *Granularity of the Dimensions*, it has been concluded that the DMMAM provides an adequate level of detail to encompass all relevant aspects of data mesh. This conclusion is supported by the presence of 54 characteristics and their definitions. The definitions were designed in a way that even the least described characteristics still possessed sufficient detail to effectively express and convey their meaning to the participants. Accurate expression and conveyance of data mesh dimensions' meaning and definitions enhance outcome reliability over multiple assessments. Fourthly, *Definition of the Measurement Attributes* pertains to the level of detail in terms of the maturity levels. Four maturity levels were selected based on the conclusion that this number would provide helpful guidance for organisations. Increasing the number of maturity levels was deemed unfavourable for reliability, as it would make it challenging for the designer, client, and user to distinguish between the maturity levels. Moreover, to ensure reliability, maturity labels and definitions were provided alongside the numerical values. Fifthly, the *Description of the Assessment Model* has been carefully considered, aiming to create an *unambiguous*, *explainable*, *self-describing*, *well-defined*, *well-articulated*, *recognisable*, and *tailored* DMMAM-design. This includes providing an introduction, instructions, control questions, criteria and requirements, and definitions. By incorporating these elements, the objectives of the assessment are better defined, and potential biases between the designer's intended meaning and the participant's interpretation are reduced. Consequently, it is expected that the results obtained from multiple assessments will exhibit consistency. Sixthly, the DMMAM aimed for *Objectivity of the Assessment Model*. Including objective measures improves the consistency of results over multiple assessments. During the design of the DMMAM, objective measures have always been the starting point. However, it was also concluded that the model should strike a balance and avoid excessive detail or complexity that could hinder its *understandability* and *client-friendliness*. Furthermore, it was determined that intangible and abstract characteristics should not be excessively objectified. Ultimately, a well-balanced set of objective and subjective measures was established.

The reliability of the assessment is evaluated by looking into the model's *Audience* and *Verification*. The *Audience* consists of organisations that are either considering or have started implementing data mesh. Furthermore, it was determined that a balanced composition of participants is crucial for a correct assessment. This balanced composition should include individuals from across the complete data supply chain, possess sufficient knowledge about all aspects of data mesh, encompass hierarchical perspectives, exhibit variation in years of working experience, and include both technical and business people. Consequently, each assessment should ideally involve twenty to thirty participants to cover all relevant perspectives. During re-assessments, the reliability of outcomes relies on maintaining a balanced group composition for each iteration. Finally, the applicability and usefulness of the DMMAM have been evaluated in cases for three organisations. The *Verification* to reliably compare the outcomes across the three organisations has been assured by having an identical assessment structure.

The validity, as defined by Creswell, refers to the extent to which the assessment results accurately represent what they are intended to measure. The validity is evaluated by looking into the objectives, constraints, and functions to assess whether the predetermined requirements have been accurately fulfilled.

According to Peffers et al. (2007), evaluating the model involves comparing the *objectives* with the observed outcomes. As a result, participants from the cases are requested to complete a questionnaire regarding the objectives of *completeness*, *orthogonality*, *explainability*, *well-definedness*, and *self-description*. The findings of the evaluation indicate that the results were predominantly positive. Specifically, *Agree* and *Strongly Agree* options were selected 65% of the time. *Disagree* was chosen in only 12.5% of cases, and *Strongly Disagree* was not selected at all. The remaining 22.5% of the results was *Neutral*. These results suggest that the objectives of the model are generally validly achieved. Among the objectives, the scores for *explainable* and *self-describing* were the most positive and had strong uniformity among participants. On the other hand, *well-defined* was the objective that generated the most disagreement among the participants.

The evaluation of the validity of the *constraints* focused on determining the extent to which the data mesh characteristics are mutually exclusive. The findings revealed that there is a large overlap among the characteristics. However, it is important to note that this overlap is considered unavoidable as data mesh involves an interplay between the dimensions. Additionally, in this research, dimension A is presented as the baseline for implementing the other dimensions, also indicating a high level of dependency. Nevertheless, future revisions of the DMMAM could strive to decrease the overlap and thereby enhance the validity. Overall, the DMMAM is still considered adequately valid for its intended purpose. The definitions of all the characteristics are formulated with the intention of outlining the different aspects, thus ensuring that participants understand the distinctions. This would decrease the probability of a participant mistakenly combining multiple characteristics into a single measure.

The assessment of the validity of the *functions* involved examining the degree to which it is appropriate to evaluate the data mesh dimensions separately, considering the strong interdependence among the dimensions and characteristics. It has been concluded that the large interplay among the dimensions would negatively affect the validity of the model outcomes for each individual dimension. Nevertheless, understanding dimensional maturity scores will benefit organisations. However, due to the overlap among characteristics, it is discouraged to evaluate only a subset of characteristics. Therefore, it is recommended to assess all characteristics to achieve accurate results.

In conclusion, the designed DMMAM is expected to be reliable and valid. However, it is important to note that this model represents the first attempt to provide a comprehensive framework for assessing data mesh maturity in organisations. Therefore, the assessment model is applicable and useful for all organisations that are considering or have initiated the implementation of data mesh and require initial guidance through a maturity assessment model. However, it is crucial to acknowledge the limitations of the model. For instance, some participants found that the assessment was not well-defined, not all criteria and requirements are objectified, and there is a large interdependence among dimensions and overlap in characteristics, which impact the reliability and validity of the DMMAM. Nevertheless, the model has been observed to be still exceedingly applicable and useful in general for its users.

9 Conclusion

This chapter presents the conclusion of this research. Due to the current lack of guidance on how to transform towards a data mesh, this research aimed to enable the assessment of how mature a data mesh implementation is within an organisation. Assessing the data mesh maturity would provide guidance to the organisation for the transformation. Therefore, the following main research question was formulated:

Main Research Question

How to assess the maturity of a data mesh implementation within an organisation?

The design of a Data Mesh Maturity Assessment Model (DMMAM) was proposed to provide guidance to organisations. The Design Science Research Methodology, provided by Peffers et al. (2007), was used to structure the design process. Main findings from the *Objective for a Solution*, *Design and Development*, *Demonstration*, and *Evaluation* phases will be presented and answers will be provided for the four sub-questions.

The *Objective for a Solution* phase aimed to explore how the DMMAM could help organisations implement data mesh. To examine the relevance of the DMMAM, the following sub-question was formulated:

Sub-Question 1

How does a data mesh maturity assessment model contribute to the data mesh implementation process?

Research conducted by De Bruin et al. (2005), García-Mireles et al. (2012), Lahrmann & Marx (2010), and Wendler (2012) demonstrated that it is acceptable to map elements from conventional maturity assessment models to data mesh, as these elements are not limited to any application domain. Consequently, reflecting the definition for a maturity assessment model provided by DAMA International (2017) on data mesh, shows that a DMMAM describes how data mesh characteristics evolve over maturity levels. As a result, Becker et al. (2009), De Bruin et al., and Mettler et al. (2010) indicated that the characteristics aligned over the maturity levels provide the transformation path towards the complete data mesh implementation. Al-Sai et al. (2022) and Król & Zdonek (2020) respectively mentioned that formulated questions and detailed descriptions, in terms of criteria and requirements, for placing a characteristic at one of the maturity levels enables the self-assessment by organisations. While self-assessing, the data mesh characteristics will be discovered and evaluated by an organisation. Moreover, the organisation will be educated about the data mesh practices, concepts, principles, and roles and responsibilities. In addition, the maturity assessment helps establish the roadmap in support of implementing data mesh. These goals from DAMA International indicated the contribution of the DMMAM to the data mesh implementation process.

The *Design and Development* phase focused on designing the DMMAM. The sub-question for developing the DMMAM was formulated as follows:

Sub-Question 2

What model could be designed to assess the maturity of a data mesh implementation within an organisation?

To design the DMMAM, the systematic engineering design approach from Dym et al. (2013) was used. The approach consists of objectives, constraints, and functional analyses to inform design choices. The objectives analysis showed that in aiming for a successful DMMAM, balance is needed between having a *valuable* model and *feasible* assessment. This trade-off between obtaining useful results while dealing with limited resources was also reflected in all constraints. This implied the number, exclusiveness, and exhaustiveness of the data mesh characteristics; the number of maturity levels; the number and background of participants in the assessment; and the duration of assessment discussion sessions. The functional analyses presented that the design needs to measure the overall data mesh maturity score; maturity scores for different data mesh dimensions; maturity scores from People, Process, Technology (PPT) perspectives; and to provide guidance for achieving higher levels of maturity.

By establishing means for the various functions and adhering to the need to enable a self-assessment, the DMMAM developed in this research consists of the following six model elements: *maturity levels*, *dimensions*, *characteristics*, *perspectives*, *questions*, and *criteria and requirements*. First of all, the data mesh maturity will be evaluated by four maturity levels, which are classified as Level 0: *Non-Initiated*, Level 1: *Conceptual*, Level 2: *Defined*, and Level 3: *Achieved*. Secondly, data mesh will be evaluated by five dimensions, which are A. *Data Foundation & Organisational Change*, B. *Domain Oriented Decentralised Data Ownership & Architecture*, C. *Data as a Product*, D. *Self-Serve Data Infrastructure as a Platform*, and E. *Federated Computational Governance*. Whereas dimensions B, C, D, and E reflect the initial data mesh dimensions introduced by Dehghani (2022a), dimension A is added as the fifth dimension to reflect on the data foundation and organisational change aspects, which need to be in place as starting point to implement the other dimensions. Thirdly, the dimensions are represented by, in total, 54 characteristics. To provide an overview, Figure 9.1 presents the DMMAM dimensions, characteristics, and levels. Fourthly, all characteristics are labelled by PPT-labels. Fifthly, all 54 characteristics have a question included to guide the participant through the assessment. At last, all maturity levels for the characteristics are described by criteria and requirements.

During the assessment, the participants will be asked to individually rate the 54 questions on the current and target level of maturity that are most fitting for the organisation. Each statement needs to be rated on a scale from Level 0:

Non-Initiated to Level 3: *Achieved*, or as *Unknown*. Conducting the assessment results in various outcomes. An overall data mesh maturity score, individual data mesh dimensional scores, and scores from PPT-perspectives will be calculated. In addition, the maturity gaps between the current and target maturity states will be obtained. The scores will be calculated by considering equal weights for the dimensions and also for all characteristics. However, the DMMAM enables the possibility to experiment with unequal weights. The strength of the different weights and the distribution of weights over the dimensions and characteristics could be adjusted manually after the assessment in consultation with the organisation. In addition, the assessment outcomes will be presented visually by radar charts and bar charts. Eventually, the various outcomes and charts will be used to benchmark organisations within or across industries.

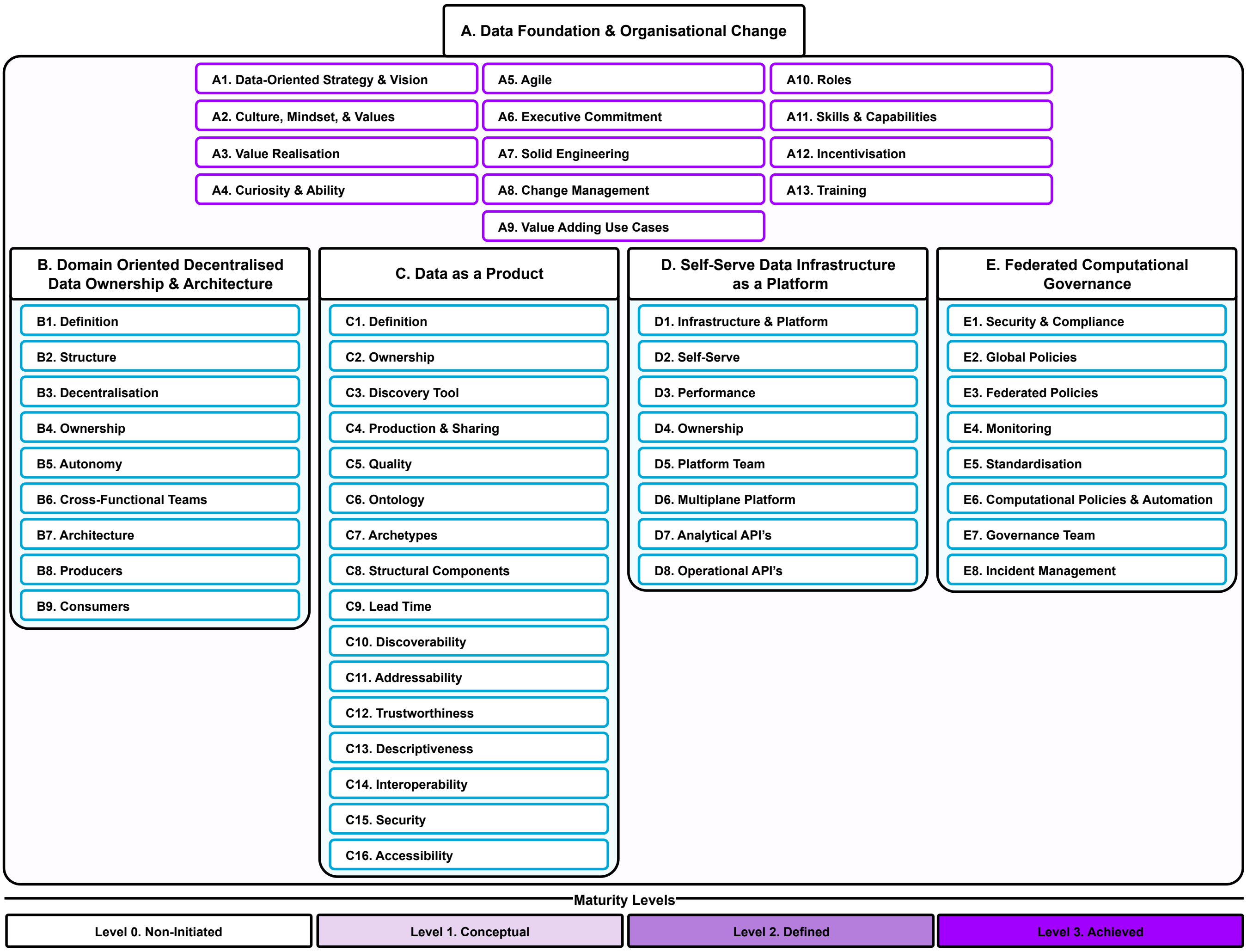


Figure 9.1: Overview Model Dimensions, Characteristics, and Levels

The applicability and usefulness of the developed DMMAM were observed during the *Demonstration* phase. Aiming for obtaining a demonstrated model, the following sub-question was formulated:

Sub-Question 3

How to demonstrate the data mesh maturity assessment model?

Three cases were carried out, involving three organisations that were individually engaged to perform a data mesh maturity assessment, with the aim of observing the applicability and usefulness of the DMMAM. The term applicability was used to assess the practical feasibility of demonstrating the DMMAM, while usefulness referred to its ability to deliver the desired outcomes in terms of the proposed maturity metrics and charts.

Three organisations, represented by a total of eight participants, participated in the data mesh maturity assessment. All participating organisations were large multinational businesses, and the participants held roles in the field of data management with 10 to 20 years of working experience. The assessment process for the organisations was carried out identically and followed three activities based on the maturity assessment structure outlined by DAMA International. These activities included planning the assessment activities, conducting the maturity assessments, and interpreting the results. The cases showed that the organisations carried out the assessments and that the model provided the defined metrics and charts. Furthermore, a comparison of outcomes was performed among the organisations, demonstrating that all three organisations attained maturity scores ranging from Level 0: *Non-Initiated* to beyond Level 1: *Conceptual*, indicating that the organisations are still in the early stages of implementing and understanding data mesh.

After performing the cases for demonstration, the DMMAM demonstrated its applicability and usefulness. The demonstration confirmed its practical feasibility for conducting a data mesh maturity assessment, establishing its applicability. Moreover, the model proved to be useful as it provided organisations with the results for the proposed maturity metrics and charts by following the selected assessment activities.

The *Evaluation* phase aimed to conclude the extent of applicability and usefulness of the DMMAM. The following sub-question was formulated to come to the final DMMAM:

Sub-Question 4

To what extent is the designed data mesh maturity assessment model applicable and useful?

To conclude on the extent of applicability and usefulness, the reliability and validity of the DMMAM were examined. Reliability, as defined by Creswell (2018), refers to the consistency of assessment results across multiple assessments conducted under the same conditions. The validity, on the other hand, is defined by Creswell as the degree to which the assessment results accurately represent what they are intended to measure.

Reliability was evaluated by discussing the evaluation criteria by Gökalp et al. (2022), which are *Fitness for Purpose*, *Completeness*, *Granularity of the Dimensions*, *Definition of the Measurement Attributes*, *Description of the Assessment Model*, *Objectivity of the Assessment*, *Audience*, and *Verification*. First of all, it has been concluded that the DMMAM fits to meet its purpose to provide guidance to organisations while implementing data mesh and that reliable results could be achieved as long as this purpose is taken into account. Secondly, the 54 characteristics represent a comprehensive and complete set of measures to evaluate the data mesh maturity. It is required that during each evaluation, all characteristics are assessed to ensure the reliability of the scores. Thirdly, the level of detail varies by the number of characteristics for each dimension and the extensiveness of the definitions. The intention was that even the least described characteristics were still sufficiently detailed to express and convey their meaning to the participants. Accurate expression and conveyance of data mesh dimensions' meaning and definitions enhance outcome reliability over multiple assessments. Fourthly, the measurement attributes were defined with four maturity levels, accompanied by labels and definitions, as it was determined that from this number the DMMAM would offer helpful guidance for organisations. Increasing the number of maturity levels was deemed unfavourable for ensuring reliability, as it would create difficulties for the designer, client, and user to distinguish between the different maturity levels. Fifthly, according to the description of the assessment model, an introduction, instructions, control questions, criteria and requirements, and definitions were provided. By incorporating these elements, potential biases arising from differences in interpretation between the designer and the participants were minimised, resulting in consistency over assessments. Sixthly, objectifying the assessment has been the starting point during the design process. However, subjective definitions and descriptions are also included to deal with value-based and abstract measures and to avoid an increase in the total number of characteristics. Eventually, a balanced set of objective and subjective measures was established. Seventhly, according to the audience, organisations which have started exploring or implementing data mesh are perceived as the target audience. Furthermore, reliability could only be assured while maintaining a balanced group in each assessment. At last, verification emphasised the importance of having a standardised case structure to enable a reliable comparison of outcomes across assessments and organisations.

Validity was evaluated by reflecting on the objectives, constraints, and functional analyses. First of all, the feedback questionnaire filled in by the case participants provided insights about the extent to which the objectives were aligned with the observed outcomes from the cases. The results showed mostly positive results, with *Agree* and *Strongly Agree* options chosen 65% of the time. *Disagree* was selected only in 12.5%, while *Strongly Disagree* was not chosen. The remaining 22.5% was *Neutral*. Secondly, the evaluation of the validity of the constraints focused on assessing the extent to which the characteristics were mutually exclusive. It has been found that there is a large overlap among the characteristics. However, it is important to note that this overlap is considered unavoidable as data mesh involves an interplay between the dimensions. However, it has been stated that the definitions for the characteristics clarified different aspects, ensuring participants understand distinctions, and avoid combining them into a single measure. Thirdly, the validity of the functions was evaluated by looking into the validity of the individual dimensional maturity scores. The overlapping characteristics negatively affect the validity of the dimensional scores. Consequently, it has been concluded that assessing only subsets of characteristics is discouraged. However, it would still be valuable for organisations to have insights into the individual dimensional scores.

In conclusion, the DMMAM has demonstrated both reliability and validity, establishing its high applicability and usefulness. However, it is important to note that this model represents the first attempt to provide a comprehensive framework for assessing data mesh maturity for organisations. Therefore, the DMMAM is applicable and useful for all organisations considering or implementing data mesh, seeking initial guidance through a maturity assessment model. However, it is crucial to acknowledge the limitations of the model. Some participants found the assessment not well-defined, various criteria and requirements are unobjectified, and there is large interdependence and overlap among dimensions. Nevertheless, the model has proven still to be exceedingly applicable and useful.

10

Discussion

Since Phase E was primarily focused on evaluating the research findings, this chapter will focus on discussing the research methodology and the research methods in Section 10.1 and 10.2 respectively.

10.1. Design Science Research Methodology

Johannesson & Perjons (2014) describe the DSRM as a method offering research structure and indicated its appropriateness for designing the DMMAM. However, this methodology also has limitations which will be discussed.

First of all, in the demonstration phase of this design science research, the applicability and usefulness of the DMMAM were observed in cases. Goecks et al. (2021) state that obtained results are highly dependent on who were involved and in which contexts the cases were performed. To illustrate, the organisations which were engaged and the participants representing the organisation were selected based on fitness and by using the professional network from Accenture. Inviting other organisations to participate would have resulted in other outcomes. Furthermore, the outcomes are also dependent on the case context. To elaborate, the design choices motivated to enable a self-assessment, having optional discussion sessions, and closing workshops. Eventually, there were no participants who requested an individual discussion session. It was also decided to conduct three cases to obtain valuable results while having a feasible assessment. In general, the people involved and the case structure affected the outcomes.

Secondly, it is needed to collect sufficient data, in terms of the number of respondents in the interviews and cases, to avoid poor quality results in design science research (Barata et al., 2022). In total, 15 experts were involved during the interviews. This number of interviewees is considered a sufficient amount. After having conducted 12 interviews, the extent of relevant information from additional interviews stabilised. Therefore, conducting more than 15 interviews would no longer improve the quality of the results in this research. To illustrate, the final interviews mainly confirmed what already was discussed during previous interviews. As a result, the formality and structuredness of the interviews evolved throughout the interview process into a more informal and open conversation. A total of eight participants were involved in the cases across three organisations. As mentioned in Section 8.1.2, this number is low compared to the aim to cover all perspectives in terms of experience and expertise.

Thirdly, Nunamaker et al. (2013) address the difficulty in producing high-impact results in the context of design science research, in particular for individual researchers. Nunamaker et al. state that making high-impact is more likely while having more collaborative research, in which multiple research methods are employed, and in which a concept is realised towards a ready-to-use deliverable. Despite this research being performed by only me as a designer, all the recommendations to increase the likelihood to make a high impact were adhered to. This research is conducted for both TU Delft and Accenture, in which many people were involved, asked for feedback, and showed interest in the final research deliverable. In addition, multiple research methods were performed, such as literature research, organising interviews, and conducting cases. At last, this research designed a DMMAM which is a deliverable which could immediately be used for real-world applications.

Fourthly, Larsen et al. (2020) mention the lack of a validity framework as an integral part of design science research. In their study, Larsen et al. introduced several validity types which should be taken into account while validating the results. The applicable validity types for this research refer to the *application*, *pragmatical*, and *semantic* validity which are defined as respectively the extent to which the designed artefact satisfies its functional needs, faithfully represents aspects of the reality, and whether the artefact used appropriate language in descriptions and components to make the model accessible for users. Due to the positive feedback throughout this research, these validity criteria are considered as satisfied.

Fifthly, Iivari (2007) addresses the ethical aspects of building IT-models as outcomes of design science research. Iivari mentions therefore the importance of constructive research, which consists of a transparent, disciplined, and rigorous process of providing design science outcomes. Constructiveness has gained large attention in this research. By using the DSRM and the systematic engineering design approach, the structure was provided. Moreover, design choices are elaborately provided, motivated, and discussed.

10.2. Research Methods

Limitations according to conducting literature research, interviews, and cases as research methods will be discussed in Section 10.2.1, 10.2.2, and 10.2.3 respectively.

10.2.1. Literature Research

Conducting literature research relies on the accessibility and availability of sources (Nakano & Muniz, 2018). Furthermore, literature research could result in the omission of relevant non-scientific publications. The authors also mentioned that it could limit creativity and intuition. In this research, the limitations of literature research were taken into account. Through literature research, the knowledge gap was explored, helped define the theoretical background, and presented best practices from existing data maturity assessment models. Despite only a few scientific publications related to data mesh being available, the book from Dehghani (2022a) served as a crucial source of knowledge to learn about data mesh. Moreover, an extensive set of Accenture internal documents was accessible during the internship to complement the scientific sources. At last, to avoid limited creativity and intuition, this research also conducted interviews and cases as research methods to obtain non-publicised knowledge.

10.2.2. Interviews

Preparing and conducting interviews are often perceived as a time-consuming process (Queirós et al., 2017). Kakilla (2021) states the potential of data loss as a limitation of conducting online interviews. In addition, online interviews could be affected by unexpected failures in the use of technology. Barriball & While (1994) state that conducting interviews with people from around the world may lead to limited exploration or questioning because of language barriers. In addition, Kakilla mentions that limited responses could be expected due to the global distinct cultural values. Due to all respondents being contacted internally within Accenture, finding 15 experts globally went effortless. Combined with the intention to directly process the interview responses towards research insights after the session, a less time-consuming process than expected was accomplished. Furthermore, all participants followed the Accenture social rules and were able to fluently speak English, which resulted in a seamless interview process. Moreover, a pilot interview was conducted to reflect on the balance of inquiry with a conversation, the duration of the interview, and the experience from the interviewee's perspective. At last, no technology issues occurred since everyone is nowadays very well acquainted with online video calling.

10.2.3. Cases

Queirós et al. present the limitation of generalising results from a small number of cases. In addition, the authors mention the potential ethical issues which could occur while presenting the results from cases in a report, in particular concerning confidentiality. Due to the limited capacity of this research, it was only possible to conduct three cases. These three cases are still considered as sufficiently valuable for demonstrating the DMMAM. However, it is not possible to make statistical generalisations from the data mesh outcomes (Teegavarapu et al., 2008; R. Yin, 1984). In addition, it was hard to make conclusions about the organisations, due to the confidentiality of participation. As a result, no clarification of why specific outcomes were obtained could be provided, such as why Organisations I and III did not fill in the target maturity states and why the assessment was not helpful for some participants.

Furthermore, Teegavarapu et al. assert that the turnaround rate, reflecting the time between the moment of sending out the questionnaire and the return of the completed form, is typically very low for questionnaires. In the cases from this research, the turnaround rate averaged over two months for all organisations, which was perceived as the slowest aspect of the entire research process. It was not expected beforehand that it would take so much time for the representative to internally find participants and distribute the questionnaire, have it completed, and returned it. This was particularly surprising, given that it had been indicated in advance that completing the assessment would not take more than only 60 minutes per person. Chapter 12 will reflect in more detail on this turnaround rate.

11

Recommendations

While reconsidering all design choices throughout this research could provide input for future research, this chapter selected five main suggestions for future research.

Ensuring the *Durability* of the Design through Ongoing Refinement and Improvement - Ongoing refinement and improvement of the designed DMMAM by data mesh SME's and data mesh practitioners will be recommended to meet the objective of having a *durable* design, defined as the ability to continue to exist for a long time, by being maintainable and sustainable. This research presents the first comprehensive framework for evaluating data mesh maturity. Moreover, the DMMAM is designed while little scientific knowledge was available and data mesh implementations at organisations were still in their infancy. As a result, the DMMAM was developed by incorporating the latest knowledge and understanding of the subject at that moment in time. Because of the limited available knowledge, various characteristics in the DMMAM were described in an abstract and subjective manner. Due to the increased enthusiasm about data mesh, it is expected more scientific and practical knowledge will become available in the future about which data mesh characteristics need to be taken into account and how these data mesh characteristics evolve over the various maturity levels. Therefore, it is suggested that the designed DMMAM needs to be frequently updated by data mesh SME's and data mesh practitioners. Frequently implies with every new reputable publication, such as a new book on data mesh, and also by monitoring new scientific publications on a quarterly to half-year basis. At the same time, this will keep the model *complete*, *unambiguous*, *measurable*, *well-articulated*, and *well-defined*.

Including Additional Guidance Would Make the Assessment More *Actionable* and *Pragmatic* - Improving the designed DMMAM by extending the provision of guidance to organisations will be suggested. The functional analysis presented and motivated the design choice to include the *Maturity gap* and *Benchmarking* means to provide guidance for achieving higher levels of maturity. Including the means *Prioritising and initiating*, *Allocating resources*, *Achievement benefits*, and *Urgency to shift* will make the DMMAM more *Actionable* and *Pragmatic*. In other words, including these means will enhance the degree to which the outcomes of the DMMAM directly translate into actionable and pragmatic next steps. How these four means would improve the guidance to organisations will be explained. First of all, prioritising and initiating the implementation of data mesh characteristics after conducting the assessment will directly drive change. For instance, the model could be extended by a classification functionality in which characteristics will be labelled according to the complexity spectrum as presented by Wells et al. (2012). The complexity spectrum by Wells et al. evaluate amongst others interventions by their intensity, the number of people involved, the needed skills, the extent of human interaction, and the clarity of its outcomes. Classification would foster the process to make quick-win interventions in a complex environment. Secondly, allocating what resources are needed to accomplish maturity progress would make the designed DMMAM more pragmatic. To illustrate, it needs to be stated how much or what budgets, training, change management programmes, up-scaling, or partnerships are required to increase levels of maturity. Thirdly, by providing the benefits of achieving higher maturity levels, organisations could motivate their employees and stakeholders to take steps forward in their development. In other words, it needs to be stated what higher maturity levels offer in terms of potential cost savings or ROI, decreases in lead time, operational efficiencies of the platform, improved compliance, and increases in user experience and satisfaction. At last, Al-Sai et al. (2022) stated that providing the design as a classification tool would offer insights into the potential risks and costs of not shifting. Presenting the urgency to increase maturity for the characteristics would make the DMMAM more actionable.

Enhancing Model Reliability and Validity by Optimal Assessment Structure - Future research into the optimal assessment structure will be suggested to improve the validity and reliability of the maturity outcomes. During the cases, the optimal research structure was not achieved due to limited resources. It would be valuable to examine the following eight aspects. First of all, to what extent would it be beneficial to have an introduction session with all participants instead of only involving the representative? Secondly, to what extent and in what form do the participants need explanation and guidance throughout the assessment process? Thirdly, would it be crucial to organise mandatory individual discussion sessions with all participants? Fourthly, what would be the exact setup during the closing workshops to create roadmaps and decide on the pace of change? Fifthly, what would be the minimum number of participants in the assessment? Sixthly, what would be the optimal composition of the group participants for having sufficient expertise according to the data mesh dimensions, balance in terms of years' experience and hierarchical perspectives, inclusion of both technical and business stakeholders, and coverage of people across the complete data supply chain? Seventhly, what would be the optimal frequency for conducting re-assessments? At last, what are the requirements in terms of assessment structure to guarantee a reliable comparison of results? In other words, which conditions need to be constant to reliably compare results over time and across organisations?

Enabling Empirical Generalisations by Improving Benchmark Functionality - Expanding the benchmarking functionality is recommended for future research to enable empirical generalisations of the maturity outcomes. Using the designed DMMAM in plenty of cases would largely increase the amount of collected data. Having sufficient data collected would enable a statistically significant comparison of the maturity outcomes across organisations within or beyond their industry boundaries. In addition, it is expected that having an extensive set of organisational maturity outcomes would strongly improve the interest of new organisations to take part in an assessment.

Examining Data Mesh as Strategy Element towards Becoming Data-Driven - Whereas the first four recommendations all focused on the further improvement of the DMMAM, this last suggestion will highlight the importance of conducting more generic research about the contribution of data mesh towards becoming a data-driven organisation. In this research, it was stated by Machado et al. (2021) that data mesh could serve as a strategy element, for organisations which experience the limitations presented by traditional data architectures, to become data-driven. However, stating that data mesh aims to overcome these limitations does not necessarily mean it will be a guaranteed success. Moreover, various challenges regarding data mesh and changing a data architecture were presented in this research. Therefore, future research is recommended to examine the extent to which the data mesh approach contributes to becoming data-driven as an organisation and would provide the presented benefits.

To help provide a starting point, it was presented by Anderson (2015) and Treder (2019) that an organisation is considered as data-driven as it has accepted the use of data at every level as a contributor to drive decision-making supported by a data culture and data processes. Reflecting on this definition from a data mesh perspective would help to evaluate the contribution. First of all, Dehghani (2022a) presents the decentralised data ownership and accountability structure in data mesh in which domains will manage the data activities autonomously. It is expected this shift could help to make data accepted at all levels of the organisation since managing data will not be perceived anymore as only an IT or Analytics task. Secondly, since data mesh is defined as a socio-technical paradigm instead of only a technical design, cultural aspects are embedded in its principles. Cultural aspects refer to the data-sharing mindset, data as a product thinking, data mesh literacy, defined values, and trust in the validity of data. As a result, this could foster the process of having a data culture in place. Thirdly, monolithic data architectures presented several drawbacks according to all data processes, such as the complex ingestion process. Data mesh tries to overcome this issue by embedding data pipelines as internal processes in the data products. Altogether, data mesh provides alternative ways to manage data where the monolithic architectures are lacking, potentially contributing towards becoming data-driven.

However, it is expected that data mesh will only be desirable for those organisations which face the limitations of monolithic data architectures. Dehghani emphasises that data mesh would have an impact on teams, ownership and accountability structures, and the delineation of responsibilities between platform, governance, and domain teams. As stated by DAMA International (2017), it is expected this change would require lots of time, effort, and investments. It is expected only organisations which experience that they get little value out of data relative to their investments, operate in complex environments, and face agility issues in the face of growth, will embrace data mesh as a solution. Nevertheless, as Driessen et al. (2023) state, it is still a long way off to argue that these assumptions are true. Future research would therefore support obtaining an understanding of the impact of data mesh towards becoming data-driven.

12 Reflection

This chapter reflects on the results of this research and the research process. Section 12.1 will reflect on the research results by looking into the societal contribution in Section 12.1.1 and scientific contribution in Section 12.1.2. In these sections, the intended societal and scientific relevance of this research, as described in Section 1.1.3 and 1.2.1 respectively, will be reflected. Section 12.2 will reflect on the research process by presenting the lessons learnt in Section 12.2.1 and by providing a personal reflection in Section 12.2.2.

12.1. Research Results

12.1.1. Societal Contribution

Section 1.1.3 presented the practical aim of this research, which is to help organisations become data-driven. To achieve this aim, this research examined the contribution of data mesh towards becoming data-driven and developed a DMMAM to guide organisations during the data mesh implementation process. Various reasons will illustrate how this aim has been achieved, after which some points of criticism will also be discussed. First of all, this research is valuable for organisations that are presently encountering limitations with monolithic data architectures, as well as those exploring, considering, or who have already started the implementation of data mesh. This research evaluated the outcomes of data mesh and suggests that it could contribute towards becoming data-driven. Secondly, this research highlighted the contribution of maturity assessment models as guidance during the implementation process. Organisations could benefit from understanding the model's elements, assessment activities, goals and drivers, and how it would offer the needed guidance. Thirdly, the DMMAM developed in this research is the first comprehensive framework for evaluating data mesh maturity, making it highly valuable for organisations. Moreover, the designed DMMAM is presented as a self-assessment which enables organisations to perform assessments in the future without the help of the designer and client. Fourthly, comparing current and target maturity scores and benchmarking them with industry competitors makes the model actionable and pragmatic after an assessment. This feature enables organisations to identify gaps and areas for improvement and to take action accordingly. Overall, it is expected organisations stand to benefit greatly from the introduction of the DMMAM. The DMMAM takes organisations a step further towards achieving the benefits of being data-driven.

However, Section 1.1.1 already mentioned the criticism on data mesh by Strengholt (2023), regarding how Dehghani (2022a) has described the approach. Strengholt raises several points in the 2nd edition of his book *Data Management at Scale*, published by O'Reilly, based on examples and experiences from over 20 years of working in various data management leadership positions at Accenture, Deloitte, ABN Amro, and Microsoft. Firstly, the approach of dividing data into domains is not suitable for restructuring large complex data landscapes that involve hundreds of application teams, thousands of services, and numerous large legacy applications to manage. Secondly, a more nuanced and pragmatic perspective on data products is needed. Advocating for a data product to be managed as a container, bundling data, metadata, code, and infrastructure together in one architecture, does not reflect how nowadays data environments operate. Thirdly, the story of data mesh is incomplete. It solely focuses on data used for analytical purposes, excluding operational purposes; it overlooks master data management; the consumer side needs to be complemented with an intelligent data fabric; and it does not provide much guidance on data modelling when building data products. Since the designed DMMAM in this research heavily relied on findings from Dehghani, this also means that this criticism is applicable to the associated characteristics, or the absence of certain characteristics, within the DMMAM. At the same time, the DMMAM could serve as a reflective tool for organisations to discover that data mesh, as described by Dehghani, also has limitations in terms of how it is presented. This would help organisations determine whether data mesh actually lives up to its expectations.

Finally, Strengholt argues that the data landscapes of future generations of organisations will be managed in completely alternative ways. This is in line with Ford et al. (2021) and Hechler et al. (2023), who suggest that the introduction of data mesh could have been expected, as it is an outstanding example of the ongoing incremental evolution observed in organisational information management. Ford et al. argue that the introduction of new capabilities brings forth new perspectives, which in turn help address enduring challenges from the past. Strengholt further states that in the coming years, organisational data architectures will be much more distributed. In short, organisations need to learn how to best balance the need for a central and decentralised approach. Both approaches have their strengths and weaknesses and it is up to the organisation to determine what is appropriate for them to implement given their organisational data landscape.

12.1.2. Scientific Contribution

Section 1.2.1 presented the contribution of this research to previously performed scientific work by combining the need for generic and concrete data mesh implementation steps including a maturity assessment. In response, the design of a DMMAM was proposed to fill this knowledge gap.

This research also followed up on previously performed data mesh research, which will be explained in more detail. Firstly, Svensson & Taghavianfar (2020) presented the benefits and challenges of becoming a data-driven organisation. This research followed up on their work by looking into the contribution of data mesh towards becoming data-driven. Secondly, this research took the suggestion for future research into account as presented by Machado et al. (2021) about a detailed approach, consisting of concrete steps, from a starting point towards the design and implementation of data mesh. Thirdly, work from Dehghani is in this research approached from the perspective of data mesh maturity. As a result, this study showed that, next to the initial four dimensions, an additional dimension is needed to appropriately assess data mesh maturity for organisations. Fourthly, Scrocca & Tommasini (2021) looked into the emergence of domain-driven data designs. This study followed up on this domain-driven data design by looking into by which characteristics the *domain oriented decentralised data ownership and architecture* dimension is reflected to measure the maturity. Fifthly, this study elaborated on the publication in which Mehmandarov et al. (2021) approached the problems from monolithic data architectures from a data mesh perspective. In addition, Mehmandarov et al. state that mapping identifiers across systems, combined with ontology-based data access, is crucial for automatic system integration. This study builds upon their research by incorporating polysemes identifiers and ontology managers in the design of DMMAM, represented respectively by the model characteristics C6: *Ontology* and C14: *Interoperability*. It has been argued that an ontology manager, standardised and provided field types, polysemes identifiers, data product global addresses, common metadata fields, schema linking, data linking, and schema stability are indicative for maturity Level 3: *Achieved*, and are therefore associated with having accomplished data product interoperability and automation. In other words, this research agrees with the statement from Mehmandarov et al. that identifiers and ontology managers are indicative of enabling automation. Sixthly, by presenting for which organisations data mesh could be beneficial as a strategy element to become data-driven, this study followed up on the article by Priebe et al. (2021) about how organisations could select their architecture paradigm. Seventhly, Joshi et al. (2021) presented in their work data governance challenges in data mesh architectures. This study followed up on examining data governance by looking into by which characteristics the *federated computational governance* dimension is reflected to measure its maturity. For them, this would open up the possibility to evaluate these characteristics in terms of potential challenges. Eighthly, Hooshmand et al. (2022) proposed a product life-cycle management (PLM) approach for transforming towards a data mesh. This research followed up on their work by more specifically examining the characteristics within the PLM-landscape, such as the *data as a product* and *self-serve data infrastructure as a platform*. Ninthly, Podlesny et al. (2022) presented that linking data products from different domains could be exploited to subvert privacy. Guaranteeing privacy within a distributed mesh has therefore its challenges. This research presented that defining security policies, governance tools, and automated processes would mitigate the risk of non-compliance according to data usage, access approval, retention, archival, and GDPR-regulations. At last, the benefits and challenges of the data mesh approach as presented by Vestues et al. (2022) are further elaborated in this research. In addition, Vestues et al. suggested future research about how a Norwegian public welfare agency, as a complex organisation, needs to implement the data mesh principles. This study provides the transformation path towards the complete data mesh implementation, which is considered helpful for them.

Overall, this research extends the current limited set of scientific publications in the field of data mesh and builds upon previous scientific work. Moreover, this research presents the first comprehensive framework for evaluating organisational data mesh maturity. In addition, this research also broadens the knowledge of data maturity assessment models. However, Sliż (2018) argues that there are too many maturity assessment models available in the literature, which raises the question about the need for additional models of this kind. Given that the maturity assessment model in this research focuses on a novel domain, it is assumed that it will provide added scientific value. At last, it is worth mentioning that only the study by Vestues et al. presented empirical findings from cases, which highlights the relevance of this research in contributing to the empirical foundation of data mesh in literature.

12.2. Research Process

12.2.1. Lessons Learnt

This section addresses the insights gained during the design of the DMMAM and the execution of the assessments. The knowledge acquired in this process will be described in order to derive lessons learnt that could contribute to the development process of new models and the future execution of maturity assessments. In addition, the usability of the model will be explained based on three suggestions on how it could be deployed from the client's perspective.

Acquired Knowledge Throughout the Design Process

Eight points of reflection will be described that have been related to, or have had an impact on, the design process of the DMMAM and the progress of the assessments.

The Form and Function of Maturity Assessment Models Have Remained Unchanged for Forty Years, with DSRM as Starting Point - The concept of the Capability Maturity Model was introduced 40 years ago to provide guidance for improving software development (Kitson & Masters, 1992). Its added value for organisations was recognised, leading to the development of numerous maturity assessment models, even over a hundred in the past few years (Adekunle et al., 2022; Becker et al., 2009). This demonstrates the demand for these models from organisations to enhance their information management processes (Steenbergen et al., 2010). This research, in which existing literature on maturity assessment models was examined, revealed that these models, regardless of their application area, exhibit uniformity in their elements and functions (Al-Sai et al., 2022; DAMA International, 2017; De Bruin et al., 2005; García-Mireles et al., 2012; Korsten et al., 2022; Lahrmann & Marx, 2010; Lasrado et al., 2015; Wendler, 2012). García-Mireles et al., Pino et al. (2008), and Staples & Niazi (2008) stated that this is mainly because these models are inspired by common standards such as CMM, ISO/IEC 15504, or CMMI-DEV during their development process (CMMI Institute, 2012; ISO, 2012; Paulk et al., 1993). This indicates that these standards would have been the formula for success in designing these models over the past forty years. Additionally, it was found by Lasrado et al. that design science research as methodology is widely used for developing maturity models.

What could be learnt from this is that when developing a maturity assessment model, the standards and methodology are essentially predetermined on how to best approach it. On one hand, it could be argued that this makes the research less experimental since the safe path is chosen. On the other hand, this approach is considered the way to go for maturity assessment model development. This raises the question of why one would deviate from it when it is the established method for developing models that are widely embraced by organisations. Consequently, organisations become familiar with the approach, making the models and their elements easily recognisable, enabling them to understand how to use them and what value could be derived from them.

The Current Absence of a DMMAM Raises the Question Why a Design for It Has Not Been Developed - This reflection point raises the question of why there is currently no maturity assessment model applied to data mesh, considering the extensive development of maturity assessment models within information management as was highlighted by Adekunle et al., Becker et al., and Steenbergen et al. Models for assessing big data, data management, data analytics, data lakes, and data warehouses already exist, as discussed in Chapter 3 by presenting the findings from SLR's conducted by Al-Sai et al., Belghith et al. (2021), and Król & Zdonek (2020). The absence of a DMMAM has elicited surprise during the identification of the knowledge gap in Section 1.2.1.

Two possible explanations are put forward for this absence by the designer. Firstly, due to the limited amount of available research on data mesh, as emphasised by Driessen et al. (2023), Goedegebuure et al. (2023), and Machado et al., and the evident successful data mesh implementations in organisations (Bode et al., 2023; Butte & Butte, 2022), it is still insufficiently understood what data mesh actually entails and how its maturity could be assessed. On the other hand, it could also be that organisations currently do not have a need for data mesh maturity assessments. The answer is likely to lie in a combination of both statements. The limited amount of literature and organisational implementations regarding data mesh do not provide guidance for defining what constitutes a successful data mesh, let alone a one-size-fits-all approach to implementation. Additionally, many organisations are still exploring data mesh and are not yet at the stage of implementation, as emphasised by Representatives P, Q, and R. At most, there are some initial initiatives within organisations to explore and experiment with the concept, but a widespread organisational transition is still a distant prospect.

The absence of a DMMAM may also provide an answer as to why completing the assessments took longer than expected in this research: it is probably still too early to evaluate the data mesh maturity for organisations. In other words, there is still insufficient incentive within organisations to gain insight into this. However, given the increasing popularity of data mesh, as stated by Goedegebuure et al. and Miner et al. (2023), it is expected that organisations will gradually adopt certain elements and move towards data mesh if it would prove the benefits beyond the limitations of traditional data architectures. It is anticipated that a few success stories from organisations will be necessary to accelerate its development. Over the years, the demand for evaluating data mesh maturity is expected to increase, and thus, the model developed in this research could serve as a solid starting point as guidance for data mesh implementation. Finally, the lesson to be learnt here is that organisations are only genuinely interested in models when they perceive value in them. Conversely, the lack of interest at present could be attributed to the absence of this model.

Discovering, Defining, and Elaborating 54 Characteristics for Measuring Data Mesh, Taking Into Account a Limited Amount of Available Research, Different Perspectives, and Time Constraints, Is an Intensive and Complex Design Process - A function of the DMMAM is to assess the overall maturity of an organisation in terms of their data mesh implementation. The aim is to indirectly measure the score of data mesh, as a latent variable, through five established dimensions, represented by numerous characteristics. Aigner et al. (1984) explain that latent variables could not be expressed as a function solely reliant on observed variables, which is also the case when measuring data mesh. Due to the limited research, as highlighted by Bode et al., Butte & Butte, Driessen et al., Goedegebuure et al., and Machado et al., it was challenging to determine which elements could be used to measure data mesh. Moreover, the characteristics themselves often possessed a latent nature, making them not directly objectively measurable through a single question. When designing the model, the designer needed to decide which elements to incorporate and how to structure the questions in a way that ensures users understand the measurements being taken, thus allowing them to provide accurate assessments. This process revealed the importance of empathising with the perspective of users and clients and dealing with limited information.

Regarding the limited availability of information, this research relied solely on scientific sources and avoided grey literature to prevent misinformation. The number of reputable sources was restricted to the book *Data Mesh* by Dehghani. However, interviews and internal documents from Accenture also offered valuable insights based on practical experiences. The deliberate choice was made to prioritise a limited variety of reputable sources over a broader range of sources with less credibility. This recommendation stems from the fact that additional scientific information could be easily incorporated in the future while removing non-academic misinformation from an existing model is more challenging.

In addition to making a latent variable measurable and handling the limited availability of information, consideration had to be given to the feasibility of the assessment, ensuring that the assessment process for the participant did not become excessively time-consuming. The initial iteration of 76 characteristics proved to be exhaustive, necessitating refinement. Definitions were formulated for all characteristics within the context of data mesh. In cases where there was a large overlap between characteristics, these were merged. Furthermore, certain characteristics were deemed inappropriate for assessing data mesh and were consequently excluded. Ultimately, this iterative process led to a reduction of 22 characteristics, resulting in a set of 54 remaining characteristics.

While establishing this set of characteristics, feedback was obtained from both users and clients. Everyone who reviewed the list argued that there is always room for improvement. However, it is crucial as the designer to make their own choices and eventually accept that a solid set of characteristics has been reached, which is sufficient for the first comprehensive version of a DMMAM. This mindset served as the foundation for the development of the DMMAM. However, during the research process, there was a recurring consideration for continuous improvement. Nevertheless, as a researcher, it is necessary to manage perfectionism within the limitations of time and resources. Moreover, definitions, questions, criteria and requirements, and labels from a PPT-perspective also needed to be established for all these characteristics. This highlights that the entire process requires significant time and effort.

During the model development, discussions were held with Representatives P, Q, and R to demonstrate the model versions to them. Due to the tight deadlines and commitments made regarding the model's delivery, pressure and responsibility were felt. Ultimately, the developed model is believed to be comprehensive and valuable for users and clients, and it was delivered within the agreed-upon timeline.

In short, measuring a latent variable such as the maturity of data mesh, with limited scientific research, and diverse opinions on what is important, necessitates striking a balance between the perspectives of users, clients, and as designer, while having little academic research to rely on. All of this, combined with the time pressure from organisations expecting an applicable and usable model, made the design process intensive and complex. The lesson to be learnt from this experience revolves around the importance of selecting a focus for gathering information, finding a balance among the different stakeholders' opinions, having confidence in an extensive and thoughtful set of characteristics that could meet the expectations of a successful model, and adhering to a realistic timeline that does not overly burden an individual designer.

An Intensive and Iterative Process of Defining Criteria and Requirements for 54 Characteristics across 4 Maturity Levels, Resulting in 216 Descriptions - The previous reflection point described the intensity and complexity of determining characteristics to measure the data mesh implementation for an organisation, given the limited available information. Additionally, the complexity also lay in the involvement of various stakeholders and the time pressure from organisations to meet agreed-upon deadlines. The same applied to establishing criteria and requirements for the four maturity levels, encompassing a total of 54 characteristics.

It was found that not all identified characteristics could be objectively measured with a single question. Subsequently, the criteria and requirements would serve as the answers to this question. Thus, it follows that it is challenging to provide answers to a question that may not have a single answer. It was important for the answers across the four maturity levels to be unidimensional, following a logical progression and distinguishing between the levels. This process was iterative, sometimes already knowing what *achieved* maturity entailed and basing the question on that. Conversely, it could also happen that the question was clear, but the four different levels still needed to be determined. This demonstrates that developing a model requires creativity from the designer, who looked for an accurate way to measure the characteristic while keeping it user-friendly and understandable for users and clients.

Furthermore, this process of establishing criteria and characteristics was carried out by an individual designer. It may have been beneficial to conduct this process in collaboration with others, for example, through structured brainstorming sessions (Dym et al., 2013). However, it was challenging to find people to collaborate with in this regard, as the process was extremely time-consuming on one hand. On the other hand, the designer was deeply immersed in the details of the design process, making it difficult for an outsider to reach the same level of detail. Given the limited resources of this research, it was better to proceed individually. Nevertheless, the input and opinions of others could have aided in achieving a better outcome.

The designer is also convinced that practical experience is necessary to determine if there is sufficient clarity for all criteria and requirements. If this is not the case for certain characteristics, revisions are necessary. The takeaway is that while theoretical considerations are important, the aim is to ensure the practical applicability and usefulness of the model for users and clients.

In short, the lesson to be drawn from this is that the process of creating 216 descriptions is an intensive and iterative process, requiring creativity to assess how a characteristic could be accurately measured while maintaining user-friendliness and understandability. Additionally, involving others could have enhanced the quality of the model, and practical applications will determine the extent to which everything is clear for users and clients.

Active Involvement of Stakeholders: Crucial for a Researcher's Reflection during the Design Process - Since this research was conducted during an internship at Accenture, and the model was intended for use by organisations, the perspectives of clients and users were explicitly incorporated into the design process. An important lesson that could be drawn is that stakeholder engagement was considered crucial for the design process of an individual researcher. This was achieved by organising interviews with the client and having informal discussion sessions with representatives of participating organisations as users. Additionally, informal conversations with experts in the field were conducted to gather insights into the research.

The consequences of not involving stakeholders in the research could potentially have resulted in an excessively lengthy and comprehensive maturity assessment model, with essential characteristics missing. Furthermore, in informal conversations, you may also receive small tips, such as the utilisation of colours in the model's colour scale. To elaborate, the colour scale, as depicted on page 3 of Appendix F, progresses from light pink to dark pink. Experts involved in the design process have informed me that this approach is preferable to using a red-to-green scale. The reason behind this preference is that the colour red could be interpreted as negative, which could create an unpleasant experience when an organisation receives a red score. Although this specific detail was not incorporated into the description of the main design process, it highlights how insights from others could be valuable in guiding one's decision-making.

In conclusion, the takeaway from this is the importance of engagement, such as incorporating the perspectives of users and clients to inform the design process and gather the necessary information.

Low Level of Participant Engagement Hampers Successful Execution of Assessments - Three cases were set up to demonstrate the model. For this purpose, three organisations were individually involved, represented by a total of eight participants. Throughout the execution of the cases, initial contact was established with a representative from each organisation who served as the point of contact. This communication remained effective during the design and development phase of the research. When the model was finally completed, a session was scheduled with the representative to explain the model and assessment structure, set expectations, and share timelines. Following the session, the questionnaire was shared with the representative, who would internally approach colleagues to complete the assessment form, taking into account the constraints regarding the minimum number of participants and diversity in expertise and experience. As a designer, I had no influence over the selection of participants who would fill in the questionnaires. I had to rely on the representative I was in contact with. Ultimately, a bottleneck became visible in this process, indicating a possible lack of engagement and willingness from others to timely complete the questionnaire. This could suggest insufficient interest and perceived value within the organisation to receive the assessment results. It is expected that there were insufficient internal incentives for the organisations to participate.

Furthermore, it is expected that involving all participants during a kick-off session would have been beneficial. This would ensure that everyone receives an explanation of the assessment's purpose, the model, the assessment process, and the intended outcomes. Additionally, participants would have an idea of who the designer is. The practice in this research demonstrated that the engagement between the designer and the participants was likely insufficient. The lesson learnt from this is that a joint kick-off is important to potentially dispel the perception among participants that it is merely a random questionnaire. It may have also been helpful to require the completion of questionnaires during a planned on-site session with the designer. This would have increased the level of engagement, and it is expected that any questions or comments from participants could have been addressed directly to the designer. In turn, this would have been useful for the designer to assess areas where the model might need further improvement.

In summary, there was a distance between the designer and the participants, which is identified as the bottleneck in the slow execution of the assessments. The lesson learnt highlights the importance of engagement and internal incentives to ensure the desired progression of the assessment.

Prioritising Representativeness over Quantity: Higher Research Capacity Would Focus on a Representative Participation rather than Increased Number of Cases - If more time had been available for this research, the preference would have been to conduct a case with twenty to thirty participants in order to enhance representativeness, rather than conducting more cases with the same assessment structure.

Conducting additional cases would have added value by allowing for the collection of sufficient data to obtain empirically generalisable results. However, given the highly intensive process of involving an organisation in a case, this is perceived as excessively time-consuming. If an organisation would have been willing to participate in a large-scale assessment, involving more than three participants, this would be of added value to the research in order to observe the influence it has on the evaluation of all aspects of data mesh. Chapter 8 indicated that twenty to thirty individuals would be necessary to adequately assess all perspectives of data mesh. However, the aforementioned reflection revealed that even engaging three participants in an assessment is challenging, let alone thirty. Nevertheless, it is expected that organising a joint kick-off session and providing sufficient internal incentives to conduct the assessment could make this feasible.

The lesson that could be drawn from this is that increasing the number of participants for representativeness is likely to be easier with an organisation that is already willing to participate, rather than approaching numerous additional organisations to set up assessments.

Continuing Model Improvement through Collaboration with an Organisation - Given the discontinuity between the designer and the participants in this study, a possible solution would be for the designer to collaborate with an organisation which is currently implementing data mesh. More specifically, this would involve the designer entering into a partnership to assist the organisation in conducting the five assessment activities from DAMA International as described in Section 3.1.3. This way, the organisation could receive guidance in implementing data mesh. Simultaneously, the organisation will actively participate in the assessment activities and will be involved in the development and improvement of the model by providing feedback. It is expected that a win-win situation could be achieved, where both parties have clear incentives to contribute.

Model Deployment Suggestions

In the previous section, it was discussed that there may currently be insufficient demand for a DMMAM. However, it is expected that due to the increased popularity of data mesh and the existence of numerous maturity assessment models, the DMMAM will also become more popular once data mesh has gained widespread adoption among organisations. The usefulness of the model will be discussed based on the findings obtained in this research.

As mentioned by De Bruin et al., conducting an assessment could, next to a self-assessment, also be facilitated by a consultancy firm. Since the goals and drivers for using maturity assessment models from the user perspective have been outlined in Chapter 3, this section considers it valuable to examine from the client's perspective how the model could be applied. Based on three suggestions, it will be explored how the model remains useful before, during, and after conducting an assessment. Firstly, the DMMAM could be used as a source of information during exploratory discussions with organisations. Organisations and consultancy firms are in constant dialogue to identify mutual opportunities, including those related to data mesh, where the consultancy firm provides advice. The developed DMMAM is considered a valuable source of information during these exploratory conversations. As mentioned by DAMA International, the model could provide insights into discovering characteristics and inform about practices, concepts, and principles, as well as identifying roles and responsibilities within data mesh. Additionally, the model could offer insights into an organisation's readiness to implement data mesh, as was stated by Al-Sai et al. and DAMA International. Moreover, the model could serve as a starting point for discussing a possible target state. Secondly, when an organisation wishes to have an assessment conducted by a consultancy firm, the project could unfold as follows. All key stakeholders from the participating organisation would be involved, and the assessment could be carried out by two consultants over a period of six weeks. It is important for the organisation to ensure the participation of twenty to thirty persons who collectively possess a balanced mix of expertise and experience. An example 6-week project timeline is provided in Figure 12.1.

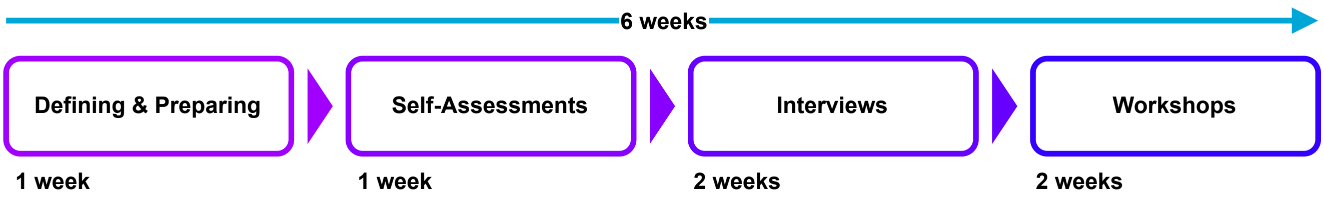


Figure 12.1: Project Timeline

In week 1, the project kick-off will take place in which the consultants, relevant stakeholders, and participants attend. In addition, the self-assessments will be shared, and the interviews and workshops will be scheduled. Week 2 focuses on conducting the self-assessment by the participants. In weeks 3 and 4, individual interviews will be organised to discuss the self-assessment with the participants. In weeks 5 and 6, the maturity scores and maturity gaps will be provided to the organisation and will be discussed. At last, these scores could be compared with industry competitors and roadmaps will be created. Thirdly, once an organisation has undergone an assessment, the implementation phase may follow, based on the established targeted programmes. Al-Sai et al. mentioned that the model could also be useful in monitoring the implementation progress, ensuring its continued utility beyond the assessment phase. To effectively track the implementation progress, DAMA International emphasises the need for periodic re-assessments.

12.2.2. Personal Experience

In this section, I will reflect more personally on the research process by highlighting six perspectives of reflection.

Creativity Through Structure: ***Design Science Research Methodology and Systematic Engineering Design Approach*** - The articles from Hevner et al. (2004), Johannesson & Perjons (2014), and Peffers et al. (2007) introduced the DSRM to me. Hevner et al. and Johannesson & Perjons focus on design science applied to information systems and technology. In their work, they state that design science aims to create artefacts in the form of systems, methods, and models that support the development of IT solutions. As a result, this seemed from the beginning appropriate for my research, in which I wanted to develop a maturity assessment model focused on data mesh. Furthermore, this methodology offered a structure that enabled the creation of a ready-to-use deliverable to apply in practice. I found the combination of research structure and a design-oriented approach appealing as it enables creativity through structure and results in a deliverable which could be used by clients and users. In addition to the DSRM, I chose to adopt the systematic engineering design approach introduced by Dym et al. during the design and development phase of this research. The systematic engineering design approach emphasises the importance of incorporating users' perspectives into the deliverable, which aligns well with the consultancy industry mindset. The systematic engineering design approach provided a structured process to create and evaluate the design of a DMMAM whose functions aimed to achieve the users' objectives while adhering to specified constraints. Furthermore, Leonard et al. (2023) characterise the design process of Dym et al. as highly rationalistic. Jonassen (2012) explains that rationalistic refers to a structured progression of analyses that ultimately yields an optimal design. It was when I examined their research closely that I realised how this structured design process reflects my approach to conducting research in general, wherein I do often strive for a rational solution. Overall, these two design approaches complemented each other well and helped me to creatively design the DMMAM while adhering to a structured process.

Master Thesis Preparation (SEN2321) Course Helped Me to Hit the Ground Running - The preparation course before starting the thesis helped me with a great head start at the beginning of my research. At the end of this course, I had almost completed my entire research proposal, and only minor changes were needed for my thesis kick-off. In addition, during the preparation course, I had ample time to find supervisors who were a good fit for my research topic and could guide me throughout the research process. Furthermore, the preparation course helped me align expectations with both TU Delft and Accenture. I planned several meetings and had email contact with my supervisors to fine-tune the research proposal. Moreover, I also created a comprehensive and feasible research plan, which was well-received by both supervisors. As a result of the thorough preparation and consultations with my supervisors, I only needed to focus on getting the research structure right after the thesis kick-off. Overall, the weeks leading up to the kick-off felt like a flying start.

The Information Gathering Process: ***Balancing Scientific Research with Accenture Internal Publications*** - Due to the novelty of data mesh, only ten scientific publications were initially found that profoundly touched upon the topic of data mesh. Despite the limited amount of available scientific work, these articles were sufficient to identify a relevant knowledge gap. At the same time, I found it interesting and exciting to research such an emerging topic. There was plenty of grey literature that demonstrated the enthusiasm around data mesh, and it felt that it was the right topic to conduct research on from the beginning. During my research, I did not focus on blogs and web articles to avoid misinformation. While exploring the topic in more depth, I came across the book titled *Data Mesh* by Dehghani, which had eventually a significant impact on my research. This book is currently regarded as the authoritative source on data mesh and provides insights directly from the founder. In addition to the scientific articles and the book, exploring all documents about data mesh published internally at Accenture during the first weeks of my internship allowed me to gain a comprehensive understanding of the topic from both a theoretical and practical perspective. Altogether, complementing scientific literature with Accenture's internal publications worked well for me to gather knowledge.

Synergising Theory and Practice: ***Complementing Academic Research with an Accenture Internship*** - The process of gathering information for my research showed that combining scientific research with Accenture's internal publications was ideal for acquiring the necessary knowledge. This synergy proved to be effective in many other ways as well. While academia provided me with relevant scientific insights about data mesh, Accenture offered an extensive set of internal publications, a global network of data management experts, and communities of data mesh enthusiasts. I was pleasantly surprised during the first few weeks of my internship to see the level of interest people showed in my research and how they wanted to stay informed throughout the process. The high level of enthusiasm for data mesh within Accenture globally was a great experience for me to be a part of. In addition, interviews were conducted to supplement the limited amount of scientific research. The global network at Accenture was ideal in finding sufficient respondents who possessed both theoretical and practical knowledge of data mesh. Accenture colleagues also helped me through professional outreach to find organisations to participate in the cases. Ultimately, this process was the most educational experience in this research for me. To illustrate, after two months of research, I began exploring which organisations would potentially be a good fit for participating in the cases. I contacted various people from different organisations and planned introduction meetings both online and in person where I presented my idea to develop a DMMAM and asked whether they would like to take part in an assessment. Convincing an organisation without being able to show a developed model at that time was a very interesting process. Gaining the trust of these organisations and presenting my progress in bi-weekly meetings was the part I enjoyed the most during this research. Due to their interest in my research, my work felt truly appreciated and I wanted to demonstrate my progress towards the final model. Finally, once my model was completed, I shared the assessment with three organisations. Despite their willingness to participate, it took much longer than expected to receive back the results, which showed the vulnerability of making oneself dependent on others' work. However, I understood that it took them longer than expected and respected their time and effort. When I received all the results, I appreciated the professionalism with which the participants filled in the self-assessments. For me, it resulted in a demonstrated DMMAM. All in all, this process of convincing organisations to take part in the assessment, keeping them involved throughout the design process, incorporating their feedback, sending out the assessments, and presenting the findings in closing workshops was the most valuable learning experience throughout my research, while realising that stakeholder management requires lots of time and effort.

Managing the Impact of Limited Organisational Capacity on Conducting Self-Assessments - Where the preparation course gave me a flying start, it turned out that carrying out the cases required more time than anticipated. This was not so much due to my capacity, but because I relied on organisations to conduct the self-assessments. The participants involved in these organisations were mostly senior managers with numerous other responsibilities to handle, indicating they would probably have many other priorities besides contributing to this research. The time it took between sending out the assessments and receiving them back averaged over two months for all organisations, even though completing the assessment itself takes only 60 minutes. Additionally, during these two months, there were often requests from my side for updates on the progress of the assessments, and assistance was offered by me in case this would be needed. However, replies were often absent or indicated that it had not been possible at that moment due to limited internal capacity. As a designer, this felt sometimes like a black box to me, with little to no knowledge of the assessment's progress or whether the results would eventually arrive. This contradicted the communication during the design and development phase, which was characterised by a high level of engagement in bi-weekly meetings. Ultimately, everything turned out fine, and the results were filled out professionally, albeit later than expected. Looking back, I could have scaled down my research in this aspect, such as by focusing on one or two organisations. However, on a personal level, this experience has facilitated my development, as this situation is often present when dealing with the involvement of large organisations and individuals who have numerous other responsibilities to manage.

Bringing People into the Data Mesh Enthusiasm: *Connecting People, Sharing Insights, and Encouraging Reflection* - The last perspective I would like to share is that I enjoyed the high level of involvement from people throughout my research. I have noticed that I am disciplined and motivated to conduct research on my own for over half a year. However, I did enjoy the interaction and collaboration with others in my research. To illustrate, I presented during several sessions about data mesh and my research results to share my data mesh enthusiasm with others. I enjoyed introducing people to the topic of data mesh and connecting data mesh enthusiasts. At the same time, involving a lot of people in your research means that everyone will have an opinion about your work. As a designer, it was important for me to make design choices by myself to keep making progress. But, having so many people involved made me feel like my work was appreciated. In total, the experts who participated in interviews, the client representatives with whom I had bi-weekly meetings during the design and development phase, the user participants who filled in the self-assessments, Accenture colleagues who helped me to get in contact with organisations, experts who assisted me with maturity assessment models or showed practical examples of data mesh implementations, various senior-level experts who helped me to evaluate the designed DMMAM, my supervisors, and lastly, the great global audience for whom I had the opportunity to present my research findings, shows a total of easily more than one hundred people who were involved throughout my internship. This is where I would like to express my gratitude to everyone at Accenture and also to my supervisors from Delft.

References

Aarons, G. A., Hurlburt, M. S., & Horwitz, S. M. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health*, *38*, 4–23. https://doi.org/10.1007/s10488-010-0327-7

Adekunle, S. A., Aigbavboa, C., Ejohwomu, O., Ikuabe, M., & Ogunbayo, B. F. (2022). A critical review of maturity model development in the digitisation era. *Buildings*, *12*(6), 858. https://doi.org/10.3390/buildings12060858

Aigner, D. J., Hsiao, C., Kapteyn, A., & Wansbeek, T. (1984). *Chapter 23: Latent variable models in econometrics* (Vol. 2). Elsevier. https://doi.org/10.1016/s1573-4412(84)02015-8

Aissi, M. E. M. E., Benjelloun, S., Loukili, Y., Lakhrissi, Y., Boushaki, A. E., Chougrad, H., & Ali, S. E. B. (2022). *Data lake versus data warehouse architecture: A comparative study* (Vol. 745). Springer. https://doi.org/10.1007/978-981-33-6893-4_19

Al-Mekhlal, M., & Khwaja, A. A. (2019). A synthesis of big data definition and characteristics. *IEEE Computer Society*, 314–322. https://doi.org/10.1109/CSE/EUC.2019.00067

Alrehamy, H. H., & Walker, C. (2018). Semlinker: Automating big data integration for casual users. *Journal of Big Data*, *5*, 14. https://doi.org/10.1186/s40537-018-0123-x

Al-Sai, Z. A., Abdullah, R., & Husin, M. H. (2019). Big data impacts and challenges: A review. *IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology*, 150–155. https://doi.org/10.1109/jeeit.2019.8717484

Al-Sai, Z. A., Husin, M. H., Abdullah, R. R., Zitar, R. A., Abualigah, L., & Gandomi, A. H. (2022). Big data maturity assessment models: A systematic literature review. *Big Data and Cognitive Computing*, *7*(1), 2. https://doi.org/10.3390/bdcc7010002

Anderson, C. (2015). *Creating a data-driven organization: Practical advice from the trenches*. O'Reilly Media. https://learning-oreilly-com.tudelft.idm.oclc.org/library/view/creating-a-data-driven/9781491916902/

Asswad, J., & Gómez, J. M. (2021). Data ownership: A survey. *Information*, *12*(11), 465. https://doi.org/10.3390/info12110465

Atwal, H. (2020). *Devops for dataops*. Apress. https://doi.org/10.1007/978-1-4842-5104-1_7

Bandari, V. (2020). Impact of data democratization and data literacy on employee productivity. *Sage Science Review of Educational Technology*, *3*(1), 37–48. https://journals.sagescience.org/index.php/ssret/article/view/42

Barata, J., da Cunha, P. R., & de Figueiredo, A. D. (2022). Self-reporting limitations in information systems design science research. *Business & Information Systems Engineering*, *65*, 143–160. https://doi.org/10.1007/s12599-022-00782-8

Barham, H., & Daim, T. U. (2020). The use of readiness assessment for big data projects. *Sustainable Cities and Society*, *60*. https://doi.org/10.1016/j.scs.2020.102233

Barriball, K. L., & While, A. (1994). Collecting data using a semi-structured interview: A discussion paper. *Journal of Advanced Nursing*, *19*(2), 328–335. https://doi.org/10.1111/j.1365-2648.1994.tb01088.x

Becker, J., Knackstedt, R., & Pöppelbuß, J. (2009). Developing maturity models for IT management. *Business & Information Systems Engineering*, *1*, 213–222. https://doi.org/10.1007/s12599-009-0044-5

Belghith, O., Skhiri, S., Zitoun, S., & Ferjaoui, S. (2021). A survey of maturity models in data management. *IEEE 12th International Conference on Mechanical and Intelligent Manufacturing Technologies*, 298–309. https://doi.org/10.1109/icmimt52186.2021.9476197

Belov, V. N., Kosenkov, A., & Nikulchev, E. (2021). Experimental characteristics study of data storage formats for data marts development within data lakes. *Applied Sciences*, *11*(18), 8651. https://doi.org/10.3390/app11188651

Berndtsson, M., Forsberg, D., Stein, D., & Svahn, T. (2018). Becoming a data-driven organisation. *26th European Conference on Information Systems: Beyond Digitization: Facets of Socio-Technical Change*. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85061304326&partnerID=40&md5=c1079b9201accc386cf1b7651f5deacc

Berndtsson, M., Lennerholt, C., Svahn, T., & Larsson, P. (2020). 13 organizations' attempts to become data-driven. *International Journal of Business Intelligence Research*, *11*(1), 1–21. https://doi.org/10.4018/IJBIR.2020010101

Blase, K. A., & Fixsen, D. L. (2013). Core intervention components: Identifying and operationalizing what makes programs work. *ASPE Research Brief*, 1–22. https://aspe.hhs.gov/reports/core-intervention-components-identifying-operationalizing-what-makes-programs-work-0#main-content

Bode, J., Kühl, N., Kreuzberger, D., & Hirschl, S. (2023). Data mesh: Motivational factors, challenges, and best practices. *arXiv*. https://doi.org/10.48550/arXiv.2302.01713

Borisyak, M., Ratnikov, F., Derkach, D., & Ustyuzhanin, A. (2017). Towards automation of data quality system for CERN CMS experiment. *Journal of Physics*, *898*. https://doi.org/10.1088/1742-6596/898/9/092041

Brennan, R., Quigley, S., De Leenheer, P., & Maldonado, A. (2018). *Automatic extraction of data governance knowledge from slack chat channels* (Vol. 11230). Springer. https://doi.org/10.1007/978-3-030-02671-4_34

Butte, V. K., & Butte, S. (2022). Enterprise data strategy: A decentralized data mesh approach. *International Conference on Data Analytics for Business and Industry*, 62–66. https://doi.org/10.1109/icdabi56818.2022.10041672

Castillo-Montoya, M. (2016). Preparing for interview research: The interview protocol refinement framework. *The Qualitative Report*, *21*(5), 811–831. https://doi.org/10.46743/2160-3715/2016.2337

Certybox Education. (2023). *Data science*. Education C. https://www.google.nl/books/edition/Data_Science/ZC2uEAAAQBAJ?hl=en&gbpv=1&dq=Data+Science+By+Certybox+Education&printsec=frontcover

Chehri, A., Saadane, R., Fofana, I., & Jeon, G. (2022). *Smart grid for sustainable cities: Strategies and pathways for energy efficiency solutions* (Vol. 263). Springer. https://doi.org/10.1007/978-981-16-6269-0_27

Chen, I. J., & Popovich, K. (2003). Understanding customer relationship management (CRM). *Business Process Management Journal*, *9*(5), 672–688. https://doi.org/10.1108/14637150310496758

Choi, M. R., Bae, J., Moon, C., & Baik, D. (2009). A methodology for developing data taxonomy for data architecture. *IEEE Fifth International Joint Conference on INC, IMS and IDC*, 833–838. https://doi.org/10.1109/ncm.2009.244

CMMI Institute. (2012). *CMMI levels of capability and performance*. Retrieved May 22, 2023, from https://cmmiinstitute.com/cmmi/intro

Correia, E., Garrido-Azevedo, S., & Carvalho, H. (2023). Supply chain sustainability: A model to assess the maturity level. *Systems*, *11*(2), 98. https://doi.org/10.3390/systems11020098

Couto, J. C., & Ruiz, D. D. (2022). An overview about data integration in data lakes. *IEEE 17th Iberian Conference on Information Systems and Technologies*, 1–7. https://doi.org/10.23919/cisti54924.2022.9820576

Creswell, J. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage. https://spada.uns.ac.id/pluginfile.php/510378/mod_resource/content/1/creswell.pdf

Cuenca, V., Urbina, M., Córdova, A., & Cuenca, E. (2021). Use and impact of big data in organizations. *Advances in Intelligent Systems and Computing*, *1327*, 147–161. https://doi.org/10.1007/978-3-030-68083-1_12

DAMA International. (1980). *About us*. Retrieved May 20, 2023, from https://www.dama.org/cpages/mission-vision-purpose-and-goals

DAMA International. (2017). *DAMA-DMBOK: Data Management Body of Knowledge* (2nd ed.). Technics Publications. https://www.dama.org/cpages/body-of-knowledge

De Bruin, T., Freeze, R. D., Kulkarni, U., & Rosemann, M. (2005). Understanding the main phases of developing a maturity assessment model. *ACIS*. https://www.researchgate.net/profile/Michael_Rosemann2/publication/27482282_Understanding_the_Main_Phases_of_Developing_a_Maturity_Assessment_Model/links/00b7d51f71c388cc54000000.pdf

De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of big data based in its essential features. *Library Review*, *65*(3), 122–135. https://doi.org/10.1108/lr-06-2015-0061

Dehghani, Z. (2019). *How to move beyond a monolithic data lake to a distributed data mesh*. Retrieved November 18, 2022, from https://martinfowler.com/articles/data-monolith-to-mesh.html

Dehghani, Z. (2022a). *Data mesh: Delivering data-driven value at scale* (1st ed.). O'Reilly. https://learning-oreilly-com.tudelft.idm.oclc.org/library/view/data-mesh/9781492092384/

Dehghani, Z. (2022b). *First generation: Data warehouse architecture*. O'Reilly. https://learning-oreilly-com.tudelft.idm.oclc.org/library/view/data-mesh/9781492092384/ch08.html#second_generation_data_lake_architectur

Dehghani, Z. (2022c). *Four principles of data mesh and their interplay*. O'Reilly. https://learning-oreilly-com.tudelft.idm.oclc.org/library/view/data-mesh/9781492092384/ch01.html#interplay_of_the_principles

Dehghani, Z. (2022d). *Operating model of data mesh principles*. O'Reilly. https://learning-oreilly-com.tudelft.idm.oclc.org/library/view/data-mesh/9781492092384/ch01.html#data_mesh_model_at_a_glance

Dehghani, Z. (2022e). *Second generation: Data lake architecture*. O'Reilly. https://learning-oreilly-com.tudelft.idm.oclc.org/library/view/data-mesh/9781492092384/ch08.html#second_generation_data_lake_architectur

Dehghani, Z. (2022f). *The shift: Data mesh dimensions of change*. O'Reilly. https://learning-oreilly-com.tudelft.idm.oclc.org/library/view/data-mesh/9781492092384/ch01.html#the_outcomes

Devlin, B. (2020). *Thirty years of data warehousing*. Retrieved May 20, 2023, from https://www.irmconnects.com/thirty-years-of-data-warehousing-part-1/

Dixon, J. (2010). *Pentaho, hadoop, and data lakes*. Retrieved May 20, 2023, from https://jamesdixon.wordpress.com/2010/10/14/pentaho-hadoop-and-data-lakes/

Driessen, S., Monsieur, G., & van den Heuvel, W. (2023). Data product metadata management: An industrial perspective. *International Conference on Service-Oriented Computing*, *13821*, 237–248. https://doi.org/10.1007/978-3-031-26507-5_19

Dym, C., Little, P., & Orwin, E. (2013). *Engineering design: A project-based introduction* (4th ed.). John Wiley & Sons. https://www.wiley.com/en-hk/Engineering+Design:+A+Project+Based+Introduction,+4th+Edition-p-9781118324585

EDM Council. (2021). *The DCAM framework*. Retrieved May 25, 2023, from https://cdn.ymaws.com/edmcouncil.org/resource/collection/AC65DC50-5687-4942-9B53-3398C887A578/DCAM_Overview_2021_update.pdf

Fahmi, N., Hidayanto, A. N., Solikin, N., Indriany, H. S., Prastya, A., & Mardiansyah, S. F. (2022). Data warehouse capability maturity model assessment for efficient monitoring process: A case study in national narcotics board. *IOP Conference Series*, *969*(1), 12055. https://doi.org/10.1088/1755-1315/969/1/012055

Faizi, R., Fkihi, S., & Afia, A. (2017). Leveraging big data to improve customer experience. *Proceedings of the 30th International Business Information Management Association Conference: Sustainable Economic development, Innovation Management, and Global Growth*, 1405–1409. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85048600444&partnerID=40&md5=c97cced314d079ce58f65badabfaded2

Farnum, M., Mohanty, L., Ashok, M., Konstant, P., Ciervo, J., Lobanov, V. S., & Agrafiotis, D. (2019). A dimensional warehouse for integrating operational data from linical trials. *Database*. https://doi.org/10.1093/database/baz039

Ferguson, L. (2004). External validity, generalizability, and knowledge utilization. *Journal of Nursing Scholarship*, *36*(1), 16–22. https://doi.org/10.1111/j.1547-5069.2004.04006.x.

Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice*, *19*(5), 531–540. https://doi.org/10.1177/1049731509335549

Fixsen, D. L., Naoom, S. F., Blase, K. A., & Friedman, R. B. (2005). Implementation research: A synthesis of the literature. *The National Implementation Research Network*. https://nirn.fpg.unc.edu/sites/nirn.fpg.unc.edu/files/resources/NIRN-MonographFull-01-2005.pdf

Ford, N., Richards, M., Sadalage, P., & Dehghani, Z. (2021). *Software architecture: The hard parts: Modern trade-off analyses for distributed architectures*. O'Reilly Media. https://www.oreilly.com/library/view/software-architecture-the/9781492086888/

Fowler, M. (2000). *Who am I?* Retrieved May 18, 2023, from https://martinfowler.com/aboutMe.html

Frick, N., Kuttner, T. F., & Schubert, P. (2013). Assessment methodology for a maturity model for interorganizational systems: The search for an assessment procedure. *46th Hawaii International Conference on System Sciences*, 274–283. https://doi.org/10.1109/hicss.2013.106

García-Mireles, G. A., Moraga, M. Á., & García, F. (2012). Development of maturity models: A systematic literature review. *16th International Conference on Evaluation & Assessment in Software Engineering*. https://doi.org/10.1049/ic.2012.0036

Giebler, C., Gröger, C., Hoos, E., Eichler, R., Schwarz, H., & Mitschang, B. (2021). The data lake architecture framework. *Gesellschaft für Informatik*, 351–370. https://doi.org/10.18420/btw2021-19

Goecks, L., De Souza, M., Librelato, T., & Trento, L. (2021). Design science research in practice: Review of applications in industrial engineering. *Gestao e Producao*, *28*(4). https://doi.org/10.1590/1806-9649-2021v28e5811

Goedegebuure, A., Kumara, I., Driessen, S., Di Nucci, D., Monsieur, G., van den Heuvel, W., & Tamburri, D. (2023). Data mesh: A systematic gray literature review. *arXiv*. https://doi.org/10.48550/arXiv.2304.01062

Gökalp, E., & Martinez, V. (2021). Digital transformation capability maturity model enabling the assessment of industrial manufacturers. *Computers in Industry*, *132*, 103522. https://doi.org/10.1016/j.compind.2021.103522

Gökalp, M., Gökalp, E., Kayabay, K., Çoban, S., Koçyiğit, A., & Eren, P. (2022). A process assessment model for big data analytics. *Computer Standards & Interfaces*, *80*, 103585. https://doi.org/10.1016/j.csi.2021.103585

Google Scholar. (2023a). *Alan Hevner*. Retrieved May 23, 2023, from https://scholar.google.com/citations?user=j9NDJ_wAAAAJ&hl=en&oi=sra

Google Scholar. (2023b). *Clive Dym*. Retrieved May 23, 2023, from https://scholar.google.com/citations?user=VBCXRSIAAAAJ&hl=en

Grossman, R. G. (2023). Ten lessons for data sharing with a data commons. *Scientific Data*, *10*(1), 120. https://doi.org/10.1038/s41597-023-02029-x

Gupta, U. (2020). *A practitioner's guide to data governance*. Emerald Insight. https://doi.org/10.1108/9781789735673

Halper, F., & Stodder, D. (2014). *TDWI analytics maturity model guide*. The Data Warehousing Institute. Retrieved May 27, 2023, from https://tdwi.org/~/media/545E06D7CE184B19B269E929B0903D0C

Hechler, E., Weihrauch, M., & Wu, Y. (2023). *Data governance in the context of data fabric and data mesh*. Apress. https://doi.org/10.1007/978-1-4842-9253-2_15

Helfert, M., & Donnellan, B. (2012). The case for design science utility: Evaluation of design science artefacts within the IT capability maturity framework. *Maynooth University*. https://mural.maynoothuniversity.ie/10825/

Hendriks, K. (2023). *Data governance structures in data mesh architectures*. University of Twente. http://essay.utwente.nl/94999/

Hevner, A., March, S., Park, J., & Ram, S. (2004). Design science in information systems research. *Management Information Systems Quarterly*, *28*(1), 75–105. https://doi.org/10.2307/25148625

Hokkanen, S. (2021). *Utilization of data mesh framework as a part of organization's data management*. University of Eastern Finland. https://erepo.uef.fi/handle/123456789/26142

Holland, S. M., Fedler, A. J., & Ditton, R. B. (1986). The group representative bias: Another look. *Leisure Sciences*, *8*(1), 79–91. https://doi.org/10.1080/01490408609513059

Holzer, M., & Karkoschka, R. (2019). *Predictive maintenance for traxon: Business model and technical solution* (Vol. 2350). Commercial Vehicles 2019. https://doi.org/10.51202/9783181023501-243

Hooshmand, Y., Resch, J., Wischnewski, P., & Patil, P. (2022). From a monolithic PLM landscape to a federated domain and data mesh. *Proceedings of the Design Society*, *2*, 713–722. https://doi.org/10.1017/pds.2022.73

Hupperz, M., Gür, I., Möller, F., & Otto, B. (2021). What is a data-driven organization? *27th Annual Americas Conference on Information Systems*. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85118666574&partnerID=40&md5=4f0529d27e74f79ccc152c35f84631db

Iivari, J. (2007). A paradigmatic analysis of information systems as a design science. *Scandinavian Journal of Information Systems*, *19*(2), 39–64. https://www.researchgate.net/publication/201168950_A_paradigmatic_analysis_of_information_systems_as_a_design_science

ISO. (2012). *ISO/IEC 15504: Information technology: Process assessment part 5: An exemplar software life cycle process assessment model*. Retrieved May 22, 2023, from https://www.iso.org/standard/60555.html

Jaiswal, A., & Bagale, P. S. (2017). A survey on big data in financial sector. *International Conference on Networking and Network Applications*, 337–340. https://doi.org/10.1109/nana.2017.46

Johannesson, P., & Perjons, E. (2014). *An introduction to design science*. Springer. https://doi.org/10.1007/978-3-319-10632-8

John Wiley & Sons. (2013). *Engineering design: A project-based introduction, 4th edition*. Retrieved May 23, 2023, from https://www.wiley.com/en-us/Engineering+Design%3A+A+Project+Based+Introduction%2C+4th+Edition-p-9781118324585

Johri, P., Kumar, A., Das, S., & Arora, S. (2017). Security framework using hadoop for big data. *International Conference on Computing, Communication, and Automation*, 268–272. https://doi.org/10.1109/ccaa.2017.8229813

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development*, *48*(4), 63–85. https://doi.org/10.1007/bf02300500

Jonassen, D. H. (2012). Designing for decision making. *Educational Technology Research and Development*, *60*(2), 341–359. https://doi.org/10.1007/s11423-011-9230-5

Joshi, D., Pratik, S., & Rao, M. (2021). Data governance in data mesh infrastructures: The saxo bank case study. *Proceedings of the International Conference on Electronic Business*, *21*, 599–604. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85125317817&partnerID=40&md5=d03c6ba8200c546167f81dacba5b5ebe

Kakilla, C. (2021). Strengths and weaknesses of semi-structured interviews in qualitative research: A critical essay. *Preprint*. https://doi.org/10.20944/preprints202106.0491.v1

Karpathiotakis, M., Floratou, A., Ozcan, F., & Ailamaki, A. (2017). No data left behind: Real-time insights from a complex data ecosystem. *Proceedings of the Symposium on Cloud Computing*, 108–120. https://doi.org/10.1145/3127479.3131208

Kerzner, H. (2001). *Strategic planning for project management using a project management maturity model*. John Wiley & Sons. http://ci.nii.ac.jp/ncid/BA53578115

Khatri, V. P., & Brown, C. L. (2010). Designing data governance. *Communications of the ACM*, *53*(1), 148–152. https://doi.org/10.1145/1629175.1629210

Khine, P. P., & Yang, X. (2018). Data lake: A new ideology in big data era. *ITM Web of Conferences*, *17*, 3025. https://doi.org/10.1051/itmconf/20181703025

Khodabandeh, A., & Palazzi, P. (1994). Software development: People, process, technology. *CERN School of Computing*. https://cds.cern.ch/record/277968/files/ecp-95-005.pdf

Kitson, D. H., & Masters, S. M. (1992). An analysis of SEI software process assessment results: 1987–1991. *International Conference on Software Engineering*, 68–77. https://resources.sei.cmu.edu/asset_files/TechnicalReport/1992_005_001_16094.pdf

Kołcz, A., Chowdhury, A., & Alspector, J. (2003). Data duplication: An imbalance problem? *ICML*. https://www.researchgate.net/publication/228547112_Data_duplication_an_imbalance_problem

Korsten, G., Aysolmaz, B., Turetken, O., Edel, D., & Ozkan, B. (2022). ADA-CMM: A capability maturity model for advanced data analytics. *Proceedings of the Annual Hawaii International Conference on System Sciences*. https://doi.org/10.24251/hicss.2022.032

Król, K., & Zdonek, D. (2020). Analytics maturity models: An overview. *Information*, *11*(3), 142. https://doi.org/10.3390/info11030142

Krosnick, J. A. (2018). *Improving question design to maximize reliability and validity*. Springer. https://doi.org/10.1007/978-3-319-54395-6_13

Krystek, M., Morzy, M., Mazurek, C., & Pukacki, J. (2023). Introducing data mesh paradigm for smart city platforms design. *Proceedings of the 56th Hawaii International Conference on System Sciences*, 6885. https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/79982307-72be-41fe-b096-529c70ef2ef0/content

Lahrmann, G., & Marx, F. (2010). *Systematization of maturity model extensions* (Vol. 6105). Springer. https://doi.org/10.1007/978-3-642-13335-0_36

Larsen, K., Lukyanenko, R., Mueller, R., Storey, V., Van der Meer, D., Parsons, J., & Hovorka, D. (2020). Validity in design science research. *Lecture Notes in Computer Science*, *12388*, 272–282. https://doi.org/10.1007/978-3-030-64823-7_25

Lasrado, L. A., Vatrapu, R., & Andersen, K. V. (2015). Maturity models development in IS research: A literature review. *Selected Papers of the IRIS*, *6*, 6. https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1005&context=iris2015

Lavasani, M., Ardali, N. R., Sotudeh-Gharebagh, R., Zarghami, R., Abonyi, J., & Mostoufi, N. (2021). Big data analytics opportunities for applications in process engineering. *Reviews in Chemical Engineering*, *39*(3), 479–511. https://doi.org/10.1515/revce-2020-0054

Lemsa, S. (2021). Challenges of advanced analytics maturity model development. *Environment. Technology. Resources.*, *2*, 88–92. https://doi.org/10.17770/etr2021vol2.6621

Leonard, A., Guanes, G., & Dringenberg, E. (2023). Undergraduate students' beliefs about diverse approaches to making engineering design decisions: Exploring change during a capstone course. *International Journal of Technology and Design Education*. https://doi.org/10.1007/s10798-022-09802-w

Lindemulder, M. (2015). *Development of a continuous improvement maturity model assessment instrument*. University of Twente. https://essay.utwente.nl/67322/1/Lindemulder_BA_BMS.pdf

Little, T. D., Bovaird, J. A., & Widaman, K. F. (2006). On the merits of orthogonalizing powered and product terms: Implications for modeling interactions among latent variables. *Structural Equation Modeling*, *13*(4), 497–519. https://doi.org/10.1207/s15328007sem1304_1

Loshin, D. (2008). *Data quality and MDM: Master Data Management*. The MK/OMG Press. https://doi.org/10.1016/B978-0-12-374225-4.X0001-X

Loshin, D. (2012). *Business intelligence: The savvy manager's guide* (2nd ed.). Morgan Kaufmann. https://doi.org/10.1016/C2010-0-67240-3

Machado, I., Costa, C., & Santos, M. (2021). Data mesh: Concepts and principles of a paradigm shift in data architectures. *Procedia Computer Science*, *196*, 263–271. https://doi.org/10.1016/j.procs.2021.12.013

Machado, I., Costa, C., & Santos, M. (2022). Advancing data architectures with data mesh implementations. *Lecture Notes in Business Information Processing*, *452*, 10–18. https://doi.org/10.1007/978-3-031-07481-3_2

Marcinkowski, B., & Gawin, B. (2021). Data-driven business model development: Insights from the facility management industry. *Journal of Facilities Management*, *19*(2), 129–149. https://doi.org/10.1108/jfm-08-2020-0051

Martins, A., Martins, P. S., Caldeira, F., & Sá, F. (2020). *An evaluation of how big-data and data warehouses improve business intelligence decision making* (Vol. 1). Springer. https://doi.org/10.1007/978-3-030-45688-7_61

McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*. Retrieved May 20, 2023, from https://hbr.org/2012/10/big-data-the-management-revolution

McCormack, K., Willems, J., Van Den Bergh, J., Deschoolmeester, D., Willaert, P., Štemberger, M. I., Škrinjar, R., Trkman, P., Ladeira, M., De Oliveira, M. R., Vukšić, V. B., & Vlahović, N. (2009). A global investigation of key turning points in business process maturity. *Business Process Management Journal*, *15*(5), 792–815. https://doi.org/10.1108/14637150910987946

Mehmandarov, R., Waaler, A., Cameron, D., Fjellheim, R., & Pettersen, T. (2021). A semantic approach to identifier management in engineering systems. *IEEE International Conference on Big Data*, 4613–4616. https://doi.org/10.1109/BigData52589.2021.9671515

Mettler, T., Rohner, P., & Winter, R. (2010). *Towards a classification of maturity models in information systems*. Physica-Verlag. https://doi.org/10.1007/978-3-7908-2404-9_39

Miller, L., Schiller, D., & Rhone, M. (2011). *Data warehouse maturity assessment service*. Teradata. Retrieved May 27, 2023, from https://assets.teradata.com/resourceCenter/downloads/WhitePapers/EB4761.pdf

Miner, G. D., Miner, L. A., & Burk, S. (2023). *The forecasts for advances in predictive and prescriptive analytics and related technologies for the year 2022 and beyond*. Elsevier. https://doi.org/10.1016/b978-0-323-95274-3.00003-8

Morris, E. (1987). Vision and strategy: A focus for the future. *Journal of Business Strategy*, *8*(2), 51–58. https://doi.org/10.1108/eb039200

Naga Rama Devi, G. (2019). Emerging trends in big data analytics: A study. *Lecture Notes in Electrical Engineering*, *500*, 563–570. https://doi.org/10.1007/978-981-13-0212-1_57

Nakano, D. N., & Muniz, J. (2018). Writing the literature review for empirical papers. *Production Journal*, *28*. https://doi.org/10.1590/0103-6513.20170086

Nambiar, A., & Mundra, D. (2022). An overview of data warehouse and data lake in modern enterprise data management. *Big Data and Cognitive Computing*, *6*(4), 132. https://doi.org/10.3390/bdcc6040132

Nunamaker, J., Twyman, N., & Giboney, J. (2013). Breaking out of the design science box: High-value impact through multidisciplinary design science programs of research. *19th Americas Conference on Information Systems: Hyperconnected World: Anything, Anywhere, Anytime*, *1*, 575–585. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84893310455&partnerID=40&md5=bde01ed8f9d5a185675315a0762c8f79

O'Reilly. (1978). *About us*. Retrieved May 18, 2023, from https://www.oreilly.com/about/

Otto, B. (2011). Organizing data governance: Findings from the telecommunications industry and consequences for large service providers. *Communications of the Association for Information Systems*, *29*. https://doi.org/10.17705/1cais.02903

Panigrahy, S., Dash, B., & Thatikonda, R. (2023). From data mess to data mesh: Solution for futuristic self-serve platforms. *International Journal of Advanced Research in Computer and Communication Engineering*, *12*(4), 677–683. https://doi.org/10.17148/ijarcce.2023.124121

Paulk, M. C., Curtis, B., Chrissis, M. B., & Cv, W. (1993). Capability maturity model, version 1.1. *IEEE Software*, *10*(4), 18–27. https://doi.org/10.1109/52.219617

Peffers, K., Tuunanen, T., Rothenberger, M., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, *24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Pino, F. J., García, F., & Piattini, M. (2008). Software process improvement in small and medium software enterprises: A systematic review. *Software Quality Journal*, *16*(2), 237–261. https://doi.org/10.1007/s11219-007-9038-z

Podlesny, N., Kayem, A., & Meinel, C. (2022). CoK: A survey of privacy challenges in relation to data meshes. *Lecture Notes in Computer Science*, *13426*, 85–102. https://doi.org/10.1007/978-3-031-12423-5_7

Pöppelbuß, J., & Röglinger, M. (2011). What makes a useful maturity model?: A framework for general design principles for maturity models and its demonstration in business process management. *Proceedings of the Nineteenth European Conference on Information Systems*. https://www.fim-rc.de/Paperbibliothek/Veroeffentlicht/327/wi-327.pdf

Priebe, T., Neumaier, S., & Markus, S. (2021). Finding your way through the jungle of big data architectures. *IEEE International Conference on Big Data*, 5994–5996. https://doi.org/10.1109/BigData52589.2021.9671862

Prodan, M., Prodan, A., & Purcărea, A. A. (2015). *Three new dimensions to people, process, technology improvement model* (Vol. 353). Springer. https://doi.org/10.1007/978-3-319-16486-1_47

Proença, D. (2016). Methods and techniques for maturity assessment. *11th Iberian Conference on Information Systems and Technologies*, 1–4. https://doi.org/10.1109/cisti.2016.7521483

Queirós, A., Faria, D., & Almeida, F. (2017). Strenghts and limitations of qualitative and quantitative research methods. *European Journal of Education Studies*, *3*(9), 369–387. https://doi.org/10.5281/zenodo.887088

Rahimzadeh, V., Lawson, J. N., Rushton, G. T., & Dove, E. S. (2022). Leveraging algorithms to improve decision-making workflows for genomic data access and management. *Biopreservation and Biobanking*, *20*(5), 429–435. https://doi.org/10.1089/bio.2022.0042

Rawat, R., & Yadav, R. (2021). Big data: Big data analysis, issues and challenges and technologies. *IOP Conference Series: Materials Science and Engineering*, *1022*(12014). https://doi.org/10.1088/1757-899X/1022/1/012014

Rivera, M., & González, K. (2022). Advantages and benefits of big data in business communication. *Smart Innovation, Systems and Technologies*, *279*, 279–292. https://doi.org/10.1007/978-981-16-9268-0_23

Santos-Neto, J. B. S. D., & Costa, A. C. (2019). Enterprise maturity models: A systematic literature review. *Enterprise Information Systems*, *13*(5), 719–769. https://doi.org/10.1080/17517575.2019.1575986

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Ashgate Publishing. https://rauterberg.employee.id.tue.nl/lecturenotes/DDM110%20CAS/Schoen-1983%20Reflective%20Practitioner.pdf

Schumacher, A., Erol, S., & Sihn, W. (2016). A maturity model for assessing industry 4.0 readiness and maturity of manufacturing enterprises. *Procedia CIRP*, *52*, 161–166. https://doi.org/10.1016/j.procir.2016.07.040

Scrocca, M., & Tommasini, R. (2021). Towards a knowledge interface for JAVA applications. *CEUR Workshop Proceedings*, *2980*. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85117734914&partnerID=40&md5=0ae880ac7b3118607c2c9cff3f79b0ed

Seawright, J., & Gerring, J. (2008). Case selection techniques in case study research. *Political Research Quarterly*, *61*(2), 294–308. https://doi.org/10.1177/1065912907313077

Shakir, A., Staegemann, D., Volk, M., Jamous, N., & Turowski, K. (2021). Towards a concept for building a big data architecture with microservices. *Business Information Systems*, *1*, 83–94. https://doi.org/10.52825/bis.v1i.67

Simitsis, A., Skiadopoulos, S., & Vassiliadis, P. (2023). The history, present, and future of ETL technology. *25th DOLAP*, *3369*. https://ceur-ws.org/Vol-3369/invited1.pdf

Sinan, I. I., Degila, J., Nwaocha, V. O., & Onashoga, S. (2022). Data architectures' evolution and protection. *International Conference on Electrical, Computer and Energy Technologies*, 1–6. https://doi.org/10.1109/icecet55527.2022.9872597

Singh, J., Singh, G., & Verma, A. (2022). The anatomy of big data: Concepts, principles and challenges. *8th International Conference on Advanced Computing and Communication Systems*, 986–990. https://doi.org/10.1109/ICACCS54159.2022.9785082

Sitarska-Buba, M., & Zygała, R. (2020). *Data lake: Strategic challenges for small and medium sized enterprises* (Vol. 887). Springer. https://doi.org/10.1007/978-3-030-40417-8_11

Sliż, P. (2018). Concept of the organization process maturity assessment. *Journal of Economics & Management*, *33*(3), 80–95. https://doi.org/10.22367/jem.2018.33.05

Soe, R., Ruohomäki, T., & Patzig, H. (2022). Urban open platform for borderless smart cities. *Applied Sciences*, *12*(2), 700. https://doi.org/10.3390/app12020700

Staples, M., & Niazi, M. (2008). Systematic review of organizational motivations for adopting CMM-based SPI. *Information & Software Technology*, *50*(7-8), 605–620. https://doi.org/10.1016/j.infsof.2007.07.003

Steenbergen, M., Bos, R., Brinkkemper, S., Van De Weerd, I., & Bekkers, W. (2010). *The design of focus area maturity models* (Vol. 6105). Springer. https://doi.org/10.1007/978-3-642-13335-0_22

Stoiber, C., Stöter, M., Englbrecht, L., Schönig, S., & Häckel, B. (2023). Keeping your maturity assessment alive. *Business & Information Systems Engineering*. https://doi.org/10.1007/s12599-023-00805-y

Strengholt, P. (2023). *Data management at scale: Modern data architecture with data mesh and data fabric* (2nd ed.). O'Reilly Media. https://learning-oreilly-com.tudelft.idm.oclc.org/library/view/data-management-at/9781098138851/

Surbakti, F. (2020). What is effective use of big data? the consensual definition of effective use of big data. *IOP Conference Series: Materials Science and Engineering*, *847*(1). https://doi.org/10.1088/1757-899X/847/1/012003

Svensson, R. B., & Taghavianfar, M. (2020). Toward becoming a data-driven organization: Challenges and benefits. *Lecture Notes in Business Information Processing*, *385*, 3–19. https://doi.org/10.1007/978-3-030-50316-1_1

Tarhan, A., Turetken, O., & Reijers, H. A. (2016). Business process maturity models: A systematic literature review. *Information & Software Technology*, *75*, 122–134. https://doi.org/10.1016/j.infsof.2016.01.010

Tariq, M., Poulin, M., & Abonamah, A. (2021). Achieving operational excellence through artificial intelligence: Driving forces and barriers. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.686624

Teegavarapu, S., Summers, J. D., & Mocko, G. M. (2008). Case study method for design research: A justification. *ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 495–503. https://doi.org/10.1115/detc2008-49980

Tiwana, A. (2014). *Platform architecture*. Elsevier. https://doi.org/10.1016/b978-0-12-408066-9.00005-9

Treder, M. (2019). *Becoming a data-driven organisation: Unlock the value of data*. Springer. https://doi.org/10.1007/978-3-662-60304-8

Vestues, K., Hanssen, G., Mikalsen, M., Buan, T., & Conboy, K. (2022). Agile data management in NAV: A case study. *Lecture Notes in Business Information Processing*, *445*, 220–235. https://doi.org/10.1007/978-3-031-08169-9_14

Voß, C. (2022). *Identifying alternatives and deciding factors for a data mesh architecture*. Hochschule für angewandte Wissenschaften München. https://dl.gi.de/server/api/core/bitstreams/97054315-86bd-418a-98a4-ed86d8bc4360/content

Wang, R. R., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, *12*(4), 5–33. https://doi.org/10.1080/07421222.1996.11518099

Webb, N., & Shavelson, R. (2005). Generalizability theory: Overview. *Encyclopedia of Statistics in Behavioral Science*, *2*, 717–719. https://doi.org/10.1002/0470013192.bsa703

Weber, K., Otto, B., & Österle, H. (2009). One size does not fit all: A contingency approach to data governance. *Journal of Data and Information Quality*, *1*(1), 1–27. https://doi.org/10.1145/1515693.1515696

Wells, M., Williams, B. G., Treweek, S., Coyle, J., & Taylor, J. (2012). Intervention description is not enough: Evidence from an in-depth multiple case study on the untold role and impact of context in randomised controlled trials of seven complex interventions. *Trials*, *13*(1), 95. https://doi.org/10.1186/1745-6215-13-95

Wendler, R. (2012). The maturity of maturity model research: A systematic mapping study. *Information & Software Technology*, *54*(12), 1317–1339. https://doi.org/10.1016/j.infsof.2012.07.007

Wider, A., Verma, S., & Akhtar, A. (2023). Decentralized data governance as part of a data mesh platform: Concepts and approaches. *IEEE International Conference on Web Services*. https://www.researchgate.net/publication/371399084_Decentralized_Data_Governance_as_Part_of_a_Data_Mesh_Platform_Concepts_and_Approaches

Wieder, P., & Nolte, H. (2022). Toward data lakes as central building blocks for data management and analysis. *Frontiers in Big Data*, *5*. https://doi.org/10.3389/fdata.2022.945720

Yin, J., Tang, Y., Lo, W., & Wu, Z. (2016). From big data to great services. *IEEE International Congress on Big Data*, 165–172. https://doi.org/10.1109/bigdatacongress.2016.28

Yin, R. (1984). *Case study research: Design and methods*. Sage. https://cds.cern.ch/record/1171670

A

Data Maturity Assessment Models

Appendix A provides an overview of existing big data, data management, and data analytics maturity assessment models presented by respectively Al-Sai et al. (2022), Belghith et al. (2021), and Król & Zdonek (2020). In these overviews are the model name, author, number of maturity levels, and number of dimensions presented. In addition, strengths and weaknesses of DMMM's will be mentioned by Belghith et al. and limitations of BDMAM's by Al-Sai et al.

A.1. Big Data Maturity Assessment Models

Table A.1 presents an overview of existing BDMAM's as presented by Al-Sai et al.

Table A.1: Big Data Maturity Assessment Models

ID	*Name*	*Author*	*Levels*	*Dimensions*
A1	TDWI Big Data Maturity Model	Halper, F.; Krishnan, K.	5	4
A2	Big Data Business Maturity Model Index	Schmarzo, B.	5	3
A3	IDC MaturityScape Big Data and Analytics	Vesset, D.; Versace, M.; Gerard, G.; O'Brien, A.; Burghard, C.; Feblowitz, J.; Osswald, D.; Ellis, S.	5	4
A4	Maturity Model for Big Data Development	van Veenstra, A.F.E.; Bakker, T.P.; Esmeijer, J.	4	9
A5	Enterprise Architecture Maturity Assessment Tool	Infotech	4	5
A6	Big Data Maturity Assessment	Knowledgent	4	5
A7	Big Data Maturity Framework	El-Darwiche, B.; Koch, V.; Meer, D.; Shehadi, R.T.; Tohme, W.	4	6
A8	Big Data Maturity Model	Radcliffe, J.	6	9
A9	A Maturity Model for Big Data and Analytics IBM	Betteridge, N.; Nott, C.	4	5
A10	Zakat Big Data Maturity Model	Sulaiman, H.; Cob, Z.C.	5	4
A11	The Big Data Temporal Maturity Model	Mach-Król, M.	5	3
A12	Hortonworks Big Data Maturity Model	Dhanuka, V.	4	5
A13	Big Data Maturity Model	Comuzzi, M.; Patel, A.	6	5
A14	A Value-Based Big Data Maturity Model	Farah, B.	5	6
A15	A Maturity Model for Big Data Analytics in Airline Network Planning	Hausladen, I.; Schosser, M.	6	4

Adapted from "Big Data Maturity Assessment Models: A Systematic Literature Review. Big Data and Cognitive Computing", by Al-Sai, Z. A., Husin, M. H., Abdullah, R., Zitar, R. A., Abualigah, L., & Gandomi, A. H., 2022, *Big Data and Cognitive Computing*, 7(1), 2. Copyright 2022 by the authors.

Table A.2 shows the limitations of existing BDMAM's as presented by Al-Sai et al. Furthermore, a ratio has been included, indicating how often a specific limitation occurred in relation to the total number of models examined.

Table A.2: Limitations of Big Data Maturity Assessment Models

Nr.	*Limitation*	*Ratio*
1	No self-assessment tool	93%
2	Assessment methods not identified	93%
3	Limited validation	93%
4	Sources of assessment components not identified	87%
5	No software assessment tool	80%
6	No visualisation report	80%
7	Poor reliability	67%
8	No evaluation in a real case study	67%
9	Development procedures not identified	67%
10	Assessment dimensions and sub-dimensions not identified	60%
11	The 5 CMM-levels not adapted	60%
12	Poor documentation about the model	33%

Adapted from "Big Data Maturity Assessment Models: A Systematic Literature Review. Big Data and Cognitive Computing", by Al-Sai, Z. A., Husin, M. H., Abdullah, R., Zitar, R. A., Abualigah, L., & Gandomi, A. H., 2022, *Big Data and Cognitive Computing*, 7(1), 2. Copyright 2022 by the authors.

A.2. Data Analytics Maturity Assessment Models

The DAMAM's examined by Król & Zdonek are presented in Table A.3.

Table A.3: Data Analytics Maturity Assessment Models

ID	*Name*	*Author*	*Levels*	*Dimensions*
B1	Analytic Processes Maturity Model	Grossman, R.L.	5	6
B2	Analytics Maturity Quotient Framework	Aryng LLC	[0, 10]	5
B3	Blast Analytics Maturity Assessment Framework	Blast Analytics & Marketing	5	6
B4	DAMM - Data Analytics Maturity Model for Associations	Association Analytics	5	4
B5	DELTA Plus Model	Davenport, T.H., Harris, J., and Morison, B.	5	7
B6	Gartner's Maturity Model for Data and Analytics	Gartner, Inc.	5	5
B7	Logi Analytics Maturity Model	Logi Analytics	5	1
B8	Online Analytics Maturity Model	Cardinal Path	5	6
B9	SAS Analytics Maturity Scorecard	SAS Institute Inc.	5	4
B10	TDWI Analytics Maturity Model	Halper, F., Stodder, D.	5	5
B11	Web Analytics Maturity Model	Hamel, S.	5	6

Adapted from "Analytics Maturity Models: An Overview", by Król, K., & Zdonek, D., 2020, *Information*, 11(3), 142. Copyright 2020 by the authors.

A.3. Data Management Maturity Models

The DMMM's as examined by Belghith et al. are presented in Table A.4.

Table A.4: Data Management Maturity Models

ID	*Name*	*Author*	*Levels*	*Dimensions*
I.	*Data Management*			
C1	DAMA-DMBOK Data Management Maturity Model	DAMA International	6	9
C2	Capability Maturity Model in Scientific Data Management	School of Information Studies Syracuse University	5	7
C3	DataFlux Master Data Management Model	DataFlux Company	5	6
C4	Research Data Management Maturity Model	School of Information Studies Syracuse University	5	5
C5	Data Management Maturity Model	CMMI Institute	5	6
C6	Data Management Capability Assessment Model	The Enterprise Data Management Council	6	7
C7	The "Orange" Data Management Maturity Model	Data Crossroad	5	4
II.	*Data/Information Governance*			
C8	IBM Data Governance Council Maturity Model	IBM	5	11
C9	DataFlux Data Governance Maturity Model	DataFlux Company	4	4
C10	The Principles Maturity Model	ARMA International	5	8
C11	Stanford Data Governance Maturity Model	Stanford University's Data Governance Office	5	6
C12	Gartner's Enterprise Information Management Maturity Model	Gartner	5	7
C13	E-ARK Information Governance Maturity Model	E-ARK	5	3
III.	*Software Development*			
C14	Capability Maturity Model	Software Engineering Institute Of Carnegie Mellon	5	3
C15	Capability Maturity Model Integration	Capability Maturity Model Institute	5	6
IV.	*Digital Assessment*			
C16	Digital Preservation Maturity Model	Preservica	5	3
C17	Digital Preservation Capability Maturity Model	Preservica	5	15
C18	Digital Assets Management Maturity Model	DAM Foundation	5	4
C19	Deloitte Digital Maturity Model	Deloitte	5	5
V.	*Analytics*			
C20.1	TDWI: Analytics Maturity Model & Assessment	Transforming Data With Intelligence	5	5
C20.2	TDWI: Self-service Analytics Maturity Model	Transforming Data With Intelligence	5	5
C20.4	TDWI: IoT Data Readiness Assessment	Transforming Data With Intelligence	5	5
C20.4	TDWI: Advanced Analytics Maturity Model	Transforming Data With Intelligence	5	5
C20.5	TDWI: Hadoop Readiness Assessment	Transforming Data With Intelligence	5	5
VI.	*Business Performance*			
C21	ECM Maturity Model	ECM	5	13
C22	Big Data Business Maturity Model Index	DELL	5	3

Adapted from "A Survey of Maturity Models in Data Management", by Belghith, O., Skhiri, S., Zitoun, S., & Ferjaoui, S., 2021, *IEEE 12th International Conference on Mechanical and Intelligent Manufacturing Technologies, 298-309*. Copyright 2021 by the authors.

Table A.5 provides various strengths and weaknesses of DMMM's as presented by Belghith et al. in their SLR.

Table A.5: Strengths and Weaknesses of Data Management Maturity Models

ID	*Family Name*	*Strengths*	*Weaknesses*
I	Data Management	Flexibility and adaptability to company specifications. Well-defined and enriched frameworks. Guidance and details on features. Best practices, actions and recommendations for maturity level evolution.	-
II	Data/Information Governance	Current state assessment. Scope definition based on priorities. Framework focuses on information governance. Risks and resources allocation.	No workshops.
III	Software Development	Recommendations for software development and data integration. Guidance for maturity improvement. Accepted and global best practices for the management and delivery of quality software processes.	Unavailable scale metrics. Require high resources and knowledge. No measurement procedures.
IV	Digital Assessment	Current state assessment for digital assets. Capabilities gaps' analysis and identification. Guidance for capabilities improvement.	No process continuity.
V	Analytics	Guidelines for all phases. Provided recommendations on future actions. Opportunity to compare results with other organisations. Opportunity to filter companies according to size or industry.	No identification of strong and weak points. Unavailable scale metrics. No training supports. The available information on other companies could be accessed and used by competitors. The limited framework focus.
VI	Business Processes	Improvement and progress tools. Guidelines for phases, processes, and business initiatives. Opportunity to compare results with other organisations.	The available information on other firms could be accessed and used by competitors.

Adapted from "A Survey of Maturity Models in Data Management", by Belghith, O., Skhiri, S., Zitoun, S., & Ferjaoui, S., 2021, *IEEE 12th International Conference on Mechanical and Intelligent Manufacturing Technologies, 298-309*. Copyright 2021 by the authors.

B

Interviews

Dym et al. (2013) present multiple ways to inform the design process. Next to examining the literature and benchmarking similar models, informal semi-structured interviews also contribute to acquiring knowledge. In this research, the main objective to conduct informal semi-structured expert interviews is to discuss the design choices of the DMMAM. In more detail, the following motivations are considered:

- Exploring the practical experience of data mesh implementations by others.
- Exploring the practical experience of performing maturity assessments by others.
- Exploring the objectives of the DMMAM.
- Exploring the constraints of the DMMAM.
- Exploring the functions of the DMMAM.
- Exploring what else needs to be taken into account while designing the DMMAM.

Appendix B will provide more background about the interview protocol in Section B.1, the interview participants in Section B.2, the interview questionnaire in Section B.3, and the interview responses in Section B.4.

B.1. Interview Protocol

Castillo-Montoya (2016) introduces a four-phase process for systematically developing an interview protocol. It will be explained how these four phases, as presented in Table B.1, are taken into account in this research.

Table B.1: Interview Protocol Framework

Phase	***Description** (Purpose)*
1	**Ensuring interview questions align with research questions.** Creating an interview protocol where interview questions are mapped against the research questions.
2	**Constructing an inquiry-based conversation.** Constructing an interview protocol that balances inquiry with conversation.
3	**Receiving feedback on interview protocols.** Obtaining feedback on interview protocol.
4	**Piloting the interview protocol.** Piloting the interview protocol in a preparation session.

Phase 1 states that interview questions need to be aligned with research questions. The interview focuses on sub-sub-question 2.2: *"What will be the design of the data mesh maturity assessment model?"* and sub-sub-question 2.3: *"What outcomes could be provided by using the data mesh maturity assessment model?"*. Phase 2 states that an interview protocol needs to balance inquiry with conversation. This balance is taken into account by writing the interview questions differently from the research questions, following the social rules of the organisation for having an ordinary conversation, including a variety of questions, and creating a script with follow-up and prompt questions. To elaborate on the variety of questions, the following four types are included: introductory questions, transition questions, key questions, and closing questions. Section B.3 will reflect on the different types while presenting the questionnaire. Phase 3 mentions the importance of asking for feedback on the interview protocol. After setting up the interview questionnaire, the questions and interview setup were discussed by the Accenture supervisor, in terms of interview structure, length, completeness of questions, comprehension, and writing style. Phase 4 states the importance of piloting the interview protocol. The interview is piloted internally at Accenture. The pilot reflected on the balance of inquiry with a conversation, the duration of the interview, and the experience from the interviewee's perspective.

B.2. Interview Respondents

Accenture as *client* was involved in the design process by conducting interviews with 15 experts who are employed at Accenture. The interviews, which lasted for 45 minutes, were conducted in an informal and semi-structured manner, allowing for a balanced approach of inquiry and conversation. The selection of these experts was based on their knowledge and demonstrated experience with data management, specifically data mesh, as well as their familiarity with maturity assessment models. All respondents are anonymised. The labels A-O are added randomly to refer to the respondents in this research. Due to potential re-identification risk, no more background about the participants than given in Table 5.2 will be provided in this research.

B.3. Interview Questionnaire

The interview is structured in four sections aligned with the variety of questions. All sections take approximately ten minutes. The questions structured an informal conversation.

Table B.2 presents the introductory questions.

Table B.2: Interview Questionnaire I

Nr.	*Question*
1	Could you please introduce yourself and describe your current role at Accenture?
2	Do you have any feedback on the questionnaire which I have to take into account during this session?
3	I contacted you since you published on the knowledge exchange about data mesh. Could you explain that specific project?
4	Next to this project, what kind of projects did you do regarding data mesh?
5	Are there specific data mesh topics wherein you are more experienced than others?
6	Have you ever conducted a maturity assessment before?
7	Which existing (data) maturity assessment models would you recommend me to evaluate as a benchmark for my model design?

Table B.3 focuses on the transition questions.

Table B.3: Interview Questionnaire II

Nr.	*Question*
8	What are the **objectives** for a data mesh maturity assessment model? Are some **objectives** more important than others? How will we know whether **objectives** have been achieved?
9	What are the **constraints** of a data mesh maturity assessment model?
10	What are the **functions** of a data mesh maturity assessment model? Given the **functions** of a data mesh maturity assessment model, what are corresponding **means**?

Table B.4 shows the key questions.

Table B.4: Interview Questionnaire III

Nr.	*Question*
11	An example **function** could be the following: *Evaluating a data mesh implementation within an organisation*. To enable this function, different data mesh dimensions need to be defined first. Which dimensions do you recommend? What are the corresponding characteristics of the different dimensions? How detailed does a maturity assessment need to be (number of characteristics per dimension)?
12	An example **function** could be the following: *Providing a final maturity score*. What score could be defined? What different levels would you suggest? How to assess the final maturity score? Do the different data mesh dimensions have different weights? Is there an overlap in the different data mesh dimensions? If so, how would you deal with this commonality?
13	An example **function** could be the following: *Providing guidance for achieving higher levels of maturity*. How could this be realised? How could the final outcomes be presented, such that it delivers relevant insights to users?

Table B.5 includes the closing questions.

Table B.5: Interview Questionnaire IV

Nr.	*Question*
14	How are the questions in a maturity assessment model aligned with the different dimensions? How many questions are needed to assess to what extent a specific characteristic is implemented?
15	Does a data mesh implementation differ across organisations from different industries?
16	What would you advise to take into consideration while designing the data mesh maturity assessment model?
17	Are there any topics we have not covered that I should take into consideration?
18	Are there any colleagues you recommend me to contact? For example, to dive more into specific aspects of data mesh?
19	Are there any final thoughts you would like to share?

B.4. Interview Responses

The findings from the interviews which seemed relevant to this research were collected. Section B.4 presents the responses for questions 8-16, including the references. Answers for the other questions are not provided due to the risk of re-identification or the answers were not contributing to this research.

Table B.6 presents the objectives and definitions.

Table B.6: Interview Results Q8

Nr.	***Objective** (Definition)*	*Reference*
1	**Valuable** Important, useful, or beneficial model for the user and client.	B
2	**Well-articulated** Able to express meanings easily and clearly and show their quality.	B
3	**Actionable** Able to be used as a reason for doing something. Outcomes need to be translated into actionable next steps.	B
4	**Tailored** To adjust or expand something to the specific needs of the user and client.	A, B, C
5	**Complete** Data mesh needs to be approached from all the different perspectives.	A, B, C
6	**Feasible** Assessment needs to be able to be performed and it needs to achieve its desired outcomes.	B
7	**Understandable** Able to be understood, so that the user and client know what something means.	D
8	**Client-friendly** Designed from the user and client's point of view. It should meet the needs of the user and client.	D
9	**Reliable** Outcome should be trusted. Important that the client could explain to the user why the final score is reliable.	C
10	**Explainable** Model and assessment outcome should be understood by the user and client.	A, C
11	**Measurable** Aspects need to be measurable to have a correct assessment.	B, C, D
12	**SMART** Characteristics need to be SMART. *Specific*: Relating to one thing and no other. *Measurable*: Able to be measured to have a correct assessment. *Achievable*: Able to be achieved. *Relevant*: Correct or suitable for a particular purpose. *Time-bounded*: Attached to a certain period.	B
13	**Pragmatic** Solving problems in a sensible way that suits the conditions that really exist now.	B
14	**Unbiased** Not affected or influenced by someone's beliefs or opinions. Maturity outcomes should be consistent regardless of who the user or client is.	B, C, D
15	**Consistent** Always behaving in a similar way. The outcome should be consistent regardless of who the user or client is.	B, C, D
16	**Self-describing** Serving to describe oneself. The user should be able to answer the questions without any help of the client.	A, C
17	**Accurate** Correct, exact, and without any mistakes. Accurate means that it is correct in all the details.	C
18	**Comfortable** Without any inconveniences for the user and client.	A
19	**Unambiguous** Expressed in a way that makes it completely clear what something means.	A
20	**Recognisable** Concepts need to be familiar to the user and client.	A
21	**Orthogonal** Independent, no overlap in dimensions and characteristics	A
22	**Well-defined** Clearly expressed, explained, and described dimensions, characteristics, and maturity levels.	A, B, I
23	**Comprehensive** Complete and including everything that is necessary.	A
24	**Modular** Consisting of separate parts that, when combined, form a complete whole.	A

Table B.7 shows the constraints and descriptions.

Table B.7: Interview Results Q9

Nr.	***Constraint*** *(Description)*	*Reference*
1	**Represented user participants must be balanced as group:** - Expertise with respect to the different data mesh dimensions. - Number of years experience. - Technical and business stakeholders. - Covering the data mesh supply chain from data producers to data consumers.	B, C, D, E, F, G, N
2	**Data mesh characteristics must be mutually exclusive and collectively exhaustive:** - All data mesh elements must be covered. - Dimensions and characteristics must be orthogonal, such that there is no overlap.	A, C
3	**Number of characteristics must balance completeness and research capacity:** - The questionnaire must be able to be completed within 60 minutes. - The questionnaire must not exceed 60 questions. - Each characteristic must have at least one question.	A, D, E, H, N
4	**Number of user participants:** - At least three user participants must be involved. - No more than six user participants must be involved.	A
5	**Duration user participant discussion session:** - Discussion session must be completed in 60 minutes.	A
6	**Number of maturity levels:** - The model must have at least four different maturity levels. - The model must have no more than five maturity levels.	A

Table B.8 includes the functions and descriptions.

Table B.8: Interview Results Q10

Nr.	***Function*** *(Description)*	*Reference*
1	**Providing maturity scores for the different data mesh dimensions.** Important to provide lower-level scores to the dimensions. Having sub-scores provide a more accurate assessment of the specific aspects. Moreover, it is clearer for the organisation where improvements could be made after the maturity assessment.	C
2	**Providing outlook by asking the target state for each characteristic.** Obtaining the gap between the current data mesh maturity state and their target maturity state.	D

Table B.9 displays the consideration regarding the data mesh dimensions.

Table B.9: Interview Results Q11 I

Nr.	***Data Mesh Dimensions***	*Reference*
1	**Four principles underpinning data mesh, by Dehghani (2022a):** I. Domain Oriented Decentralised Data Ownership and Architecture II. Data as a Product III. Self-serve Data Infrastructure as a Platform IV. Federated Computational Governance	A, B, C, G, J, K, M, N
2	**People, Progress, Technology:** Including the golden triangle.	A, B, L

Table B.10 provides the characteristics with regard to the data mesh dimension: *Domain Oriented Decentralised Data Ownership and Architecture*.

Table B.10: Interview Results Q11 II

Nr.	***Characteristics Domain Oriented Decentralised Data Ownership and Architecture***	*Reference*
1	**Data Domain Owner** Defined domain owner for each data mesh domain. The domain owner is accountable and responsible for the governance of the domain.	C, M, O
2	**Defined Data Domains** Number of domains onboarded in the data mesh.	G, L
3	**Domain Responsibility** Educating the domains and enable them to take ownership and responsibility.	K

Table B.11 provides the characteristics with regard to the data mesh dimension: *Data as a Product*.

Table B.11: Interview Results Q11 III

Nr.	***Characteristics Data as a Product***	*Reference*
1	**Data Product Sharing Incentives** How incentives are incorporated in the organisation for data product sharing. How advanced the incentivisation mechanism is to share and disclose the data products for the producers.	I, K, L, N
2	**Useful Data Products: *Usability Attributes*** Secure, Discoverable, Addressable, Understandable, Truthful (Trustworthy), Natively Accessible, Interoperable, and Valuable.	K, M
3	**Data Ontology** Expresses a relationship between two entities.	I, M
4	**Data Life-cycle** Domain team is responsible for the operations of the data product during its entire life-cycle.	M
5	**Reusable Data Products** Create a mesh of reusable data products that could be created once and shared across multiple analytical systems and workloads.	G
6	**Number of Data Products**	G, L
7	**Number of Changes in Data Products**	G
8	**Data Product Quality** Quality Levels: Bronze, Silver, and Gold.	G, L
9	**Acceptance Data as a Product** Acceptance across the domains and data supply chain.	G
10	**Rewarding Data Products** Budget to incentivise data product sharing.	G, I, K
11	**Data Product Thinking** Data as a product is about applying product thinking to how data is modelled and shared.	I

Table B.12 provides the characteristics with regard to the data mesh dimension: *Self-Serve Data Infrastructure as a Platform*.

Table B.12: Interview Results Q11 IV

Nr.	***Characteristics Self-serve Data Infrastructure as a Platform***	*Reference*
1	**Data Sharing as a Service**	M
2	**Cloud (AWS, Azure, SAP) Data Science Platform (Domain Connection)** Enabling data products sharing over different domains.	C, G, I
3	**Number of People Working on the Platform** Usage Rate.	A, G
4	**Number of Connected Systems or Sources**	G
5	**Number of Users Interacting via Marketplace**	G, L

Table B.13 provides the characteristics with regard to the data mesh dimension: *Federated Computational Governance*.

Table B.13: Interview Results Q11 V

Nr.	***Characteristics Federated Computational Governance***	*Reference*
1	**Data Lineage** Non-automated to automated process.	B, L, O
2	**Data Catalogue** Non-automated to an automated process where data products could be found. Data product owners are mentioned, which roles and responsibilities. Additionally, request data products and shop.	B, M, O
3	**Embedded Governance Rules** The extent to which governance rules are embedded in the organisation. For example, what happens in terms of data ownership to data products which are made from other data products, so-called derived data products.	L, N, O
4	**Automated Governance Rules** To what extent documentation is an updated mechanism.	N, O
5	**Documentation Responsibilities Data Producer** Ownership, legal, or compliance responsibilities.	I, N, O
6	**GDPR-Compliance** GDPR-regulations and agreements.	M
7	**Permission Allocation** Having the right permissions for the right people to obtain data.	M
8	**Data or Product Thinking/Community** Willingness to share the data knowledge.	L, M, O
9	**Change Management** Organisational aspects of data mesh: The formation of a federated governance operating model, formation of cross-functional - business, dev, data, ops-teams, and establishment of the data product ownership roles.	G, O
10	**CDO-Office** Many different responsibilities that require data specialisation fall under the CDO's functional organisation.	C
11	**Data Security** Confidential information and privacy governance rules.	C, I, K
12	**Standardisation of Security Rules** Standardisation across different domains.	C
13	**Defined Data Mesh Functions/Roles** Define clear roles for the people who perform tasks within the data mesh. Traditional: Data owner (accountable high level), data steward (operational accountable from a business perspective), data custodian (operational accountable from IT perspective). Data mesh adds: a domain owner, data product owner, and data mesh platform product owner.	C, I, O
14	**Decentralised Domain Governance** There is a decentralised governance in place. Every domain is responsible for its governance.	O
15	**Data Product Governance** Embedded governance at each data product. The domain itself defines it for the data product.	O
16	**Central Overarching Data Governance**	O

Table B.14 presents the consideration of the number of maturity levels.

Table B.14: Interview Results Q12 I

Nr.	*Number of Maturity Levels*	*Reference*
1	**Four Levels** 0-3 (0: nothing implemented, 3: completely implemented).	A, H
2	**Five Levels** 0-4, or 1-5. Easy to have two extremes (1 and 5), a middle level (3) and two levels in between (2 and 4).	A, E

Table B.15 shows the maturity score weighting options.

Table B.15: Interview Results Q12 II

Nr.	*Maturity Score Weighting*	*Reference*
1	**Equal Weights**	D, H
2	**Different Weights** Including critical and non-critical characteristics (red-flag approach).	D

Table B.16 presents means to provide guidance and recommendations for achieving higher maturity levels.

Table B.16: Interview Results Q13

Nr.	*Providing Guidance for Achieving Higher Levels of Maturity*	*Reference*
1	**Maturity Gap** Include target levels next to the current maturity levels. This opens the possibility for roadmaps.	D, E, F, H
2	**Prioritising and Initiating** Next steps are an important part of the maturity assessment. Otherwise, the maturity assessment will only be just a number. Maturity assessment should really change the process. The maturity assessment should give the opportunity to prioritise and initiate the increase of maturity and roadmap them.	A, B
3	**Allocating Resources** Budget (Generic top-down executive sponsorship), Training, Change Management Programmes, Up-scaling, Effort, Partnering with clients.	E, M, O
4	**Benchmarking** Maturity assessment model as a benchmark for competitors in the market.	A
5	**Achievement Benefits** Providing insights about what higher maturity levels concerning a specific characteristic could offer for the organisation. A description of the levels should tell you about the gaps between the maturity levels.	A
6	**Urgency to Shift** Show the urgency of shifting.	A

Table B.17 discusses the alignment between characteristics and questions.

Table B.17: Interview Results Q14

Nr.	*Alignment Characteristics & Questions*	*Reference*
1	**One-On-One Approach** Questionnaire must not exceed 30-60 questions and must be completed in a discussion in 60 minutes. Approximately one minute is required per question to complete. Each characteristic must have at least one question.	A, D, E, H, N

Table B.18 discusses the potential difference of data mesh across industries.

Table B.18: Interview Results Q15

Nr.	*Data Mesh Across Industries*	*Reference*
1	**Data mesh is Not Different across Industries, Data Mesh Maturity is Different.** It is argued that the implementation of data mesh does not differ between sectors, as data mesh stands on the same dimensions regardless of the industry. However, it is acknowledged that some sectors are further along in the implementation of data mesh.	B, D, I, N

Table B.19 provides final advice to take into consideration while designing the data mesh maturity assessment model.

Table B.19: Interview Results Q16

Nr.	*Advice*	*Reference*
1	**Leave Space for Comments** It is recommended to add a comments/notes section next to the questions. To enable the participant to easily write down his/her remarks.	D

C

Initial Set of Characteristics

This appendix provides in Table C.1 the initial set of 76 characteristics found in literature and by performing interviews.

Table C.1: Initial Set of Characteristics

ID	*Characteristic*	*Reference*	*ID*	*Characteristic*	*Reference*
A1	Data-Oriented Strategy	Z	C11	Ontology	I, M, Z
A2	Offering Value by Data & Analytics	Z	C12	Structural Components	Z
A3	Organisation Curiosity	Z	C13	Discoverability	K, M, Z
A4	Vision	Z	C14	Addressability	K, M, Z
A5	CDO & Executive Commitment	C, Z	C15	Trustworthiness	K, M, Z
A6	Change Management	G, O, Z	C16	Descriptiveness	K, M, Z
A7	Value Adding Use Cases	Z	C17	Interoperability	K, M, Z
A8	Skills & Capabilities	Z	C18	User Confidence & Trust	Z
A9	Roles	C, I, O, Z	C19	User Satisfaction	Z
A10	Rewarding & Incentivisation	G, I, K, Z	C20	Security & Accessibility	C, I, K, M, Z
A11	Change	Z	C21	Schema Evolution	Z
A12	Culture & Mindset	I, L, M, O, Z	D1	Data Source Onboarding	G, Z
A13	Democratisation	Z	D2	Self-Service Discovery	Z
A14	Training	Z	D3	Cloud Platform	C, G, I
A15	Values	Z	D4	Architecture	Z
A16	DevOps	Z	D5	Self-Serve Platform Team	Z
A17	DataOps	Z	D6	API & Protocols	Z
A18	MLOps	Z	D7	Multiplane Platform Architecture	Z
A19	Performance Tracking	Z	D8	Domain Analytical Data Interfaces	Z
A20	Awareness Importance	G, Z	D9	Domain Operational Data Interfaces	Z
A21	Incident Management	Z	D10	Legacy System Integration	Z
B1	Definition	Z	E1	Security	C, I, K, M, Z
B2	Structure	G, L, Z	E2	Global SLO	O, Z
B3	Owner Allocation	C, M, O, Z	E3	Maintenance	Z
B4	Autonomy	K, Z	E4	Compliance	M, Z
B5	Formation Cross-Functional Teams	Z	E5	Governance Adoption	Z
B6	Producers	Z	E6	Policy Coverage	Z
B7	Consumers	Z	E7	Fitness Functions	Z
C1	Definition	Z	E8	Data Governance Monitoring	Z
C2	Ownership Assignment	I, N, O, Z	E9	Standardisation for Interoperability	Z
C3	Embedded Governance	L, N, O, Z	E10	Computational Decision-Making	Z
C4	Discovery Tool	B, M, O, Z	E11	Federated Governance	O, Z
C5	Sharing	I, K, L, M, N, Z	E12	Governance Team	Z
C6	Publication	G, L, Z	E13	Domain Governance Lead	Z
C7	Lineage	B, L, O, Z	E14	Data Governance Automation	N, O, Z
C8	Marketplace	G, L	E15	Data Product Sharing Approval Process	Z
C9	Quality	G, L, Z	E16	Data Taxonomy & Glossary	Z
C10	Metadata	Z	E17	Federated Governance Operating Model	Z

D

Criteria and Requirements

Section 5.6.3 presented the characteristics that need to be assessed in the questionnaire. The process of developing the criteria and requirements will be described in this appendix in more detail.

The process of establishing the 54 characteristics involved reading the book from Dehghani (2022a) and conducting interviews. In the process of determining all the criteria and requirements, internal documents from Accenture were also examined. For all 54 characteristics, information was gathered, encompassing any relevant details that could contribute to describing each characteristic. In an iterative manner, descriptions were created in draft versions for all characteristics. More precisely, everything that could be important for a characteristic from the perspective of measuring its maturity was documented. This includes sub-elements, definitions, or how a particular progress or implementation would look in practice. It also encompasses all activities, tools, standards, and people or resources related to this. It was noticed that the descriptions of some characteristics were more uniform, while others possessed a wide range of diverse aspects. As a result, the descriptions often became quite extensive. This will be illustrated through the examination of three random characteristics, namely B4. Domain *Ownership*, C14. *Interoperability*, and E3. *Federated Policies*. Where B4 had a reasonably short description, E3 has a more extensive description, and C14 has an average description size. The draft versions of the descriptions were as follows:

B4. Domain *Ownership* - *Responsibility and accountability for domains are decentralised. Domain owners have the resources to act and drive innovation. Mesh-wide decisions are made through consensus among domain owners. Ownership is well-defined and established. The emphasis is placed on the data shared within each domain.*

C14. Interoperability - *It works in conjunction with other data products, utilising a global address system. The system includes common metadata fields, field types, and identifiers. The catalogue defines data key relationships, making it easy to understand how different data sets relate to each other. Considerations for joinability are designed to be easily comprehensible. The portfolio of data products is designed to minimise duplication. An ontology could be employed to represent standardised business domains and their relationships, enabling the creation of a globally interoperable graph of data products. Several metrics could be used to evaluate the platform, including the coverage of encoded and adopted platform policies, the version of the platform used by data products, and the ratio of data products utilising the platform compared to others. To ensure interoperability, organisations should standardise the following: field types by implementing a common, explicitly defined type system; identifiers for polysemes entities* *that could be universally recognised between different data products; global addresses for data products, using a uniform scheme for establishing linkages among the data products; common metadata fields that include time representation for when data occurs and when it is recorded; schema linking, allowing for the linking and reuse of schemas defined by other data products; data linking, enabling the mapping or linking to data in other data products; and schema stability, ensuring an approach to evolving schemas that maintains backward compatibility.*

E3. Federated Policies - *Number of domains participating in the federated governance operation. The governance operation establishes the principles that guide the decision-making process and determines what policies the organisation must implement globally and what could be left to domains. The ratio of domains and data product owners who are active members of the global federated governance. The ratio of data products implementing the latest version of policies. Data mesh governance delegates the responsibility of modelling and ensuring the quality of the data to individual domains and heavily automates the computational instructions that ensure data is secure, compliant, of quality, and usable. A federated model with computational policies embedded in the nodes of the mesh. Striving to implement automated localised capabilities at scale, using the best domain-specific technology available, which adheres to global standards. A federated organisation with a global team and localised domain teams. Federated custodianship of data by domains. Global responsibility for modelling data elements that cross multiple domain boundaries and local responsibility for domain-specific data mesh. Governance that enables autonomy. Balancing between establishing central governance to promote the re-usability of data products and maintaining autonomy in the domains. Domain-specific policies and standards. A federated team of domain representatives. Global policies are automated by the platform. Delegating governance responsibilities to autonomous domains and their data product owners. Granting domain teams the autonomy to move fast independently.*

In a similar fashion, this was done for all 54 characteristics. Once all the descriptions were drafted, an attempt was made to extract the essence from each description, considering how it could also be aligned with the maturity scale to assess the extent of its implementation. Ultimately, the core of each description, determined by the designer, became the definition of the characteristic. The descriptions for B4, C14, and E3 were established as follows:

B4. Domain *Ownership* - *Defined and allocated domain owner.*

C14. Interoperability - *Enabled ability to correlate data products across domains and stitch them together (join, filter, aggregate).*

E3. Federated Policies - *Adopted federated policies and standards by domains that guide the decision-making process and decide about what policies the organisation must implement globally and what could be left to domains.*

Based on this definition, the next step was to determine how the classification could be made in relation to the four levels of maturity. Firstly, Level 0 and Level 3 were filled in, as they represent the states of *Non-Initiated* and *Achieved* respectively. Simultaneously, careful consideration had to be given to how the question needs to be formulated, as the descriptions of the different maturity levels for a characteristic provide the multiple-choice answers to this question. This highlights the iterative nature of the process. While setting up the descriptions for the criteria and requirements, the activities, tools, standards, and people or resources related to the characteristic were looked into again. Furthermore, the definitions of the various maturity levels, as presented in Table 5.8, were taken into account when creating the categorisation of the different levels. Once the descriptions for Level 0 and Level 3 were established, Level 1: *Conceptual* and Level 2: *Defined* were also filled in. This process was carried out for all 54 characteristics, which indicates the tremendous amount of work it has required.

E

Aggregated Responses Cases

Appendix E presents the aggregated responses from the cases for Organisation I, II, and III in Table E.1, E.2, and E.3 respectively. The responses provided are the current maturity level, target maturity level, and maturity gap.

Table E.1: Aggregated Responses Organisation I: Characteristics

ID	*Characteristic*	*Current Maturity Level*	*Target Maturity Level*	*Maturity Gap*
A1	Data-Oriented Strategy & Vision	1.00	NA	NA
A2	Culture, Mindset, & Values	0.67	NA	NA
A3	Value Realisation	0.67	NA	NA
A4	Curiosity & Ability	1.00	NA	NA
A5	Agile	0.33	NA	NA
A6	Executive Commitment	0.33	NA	NA
A7	Solid Engineering	1.00	NA	NA
A8	Change Management	0.33	NA	NA
A9	Value Adding Use Cases	0.33	NA	NA
A10	Roles	1.00	NA	NA
A11	Skills & Capabilities	0.33	NA	NA
A12	Incentivisation	0.00	NA	NA
A13	Training	0.00	NA	NA
B1	Definition	1.00	NA	NA
B2	Structure	1.00	NA	NA
B3	Decentralisation	1.33	NA	NA
B4	Ownership	0.33	NA	NA
B5	Autonomy	0.67	NA	NA
B6	Cross-Functional Teams	0.67	NA	NA
B7	Architecture	0.67	NA	NA
B8	Producers	0.33	NA	NA
B9	Consumers	0.33	NA	NA
C1	Definition	0.00	NA	NA
C2	Ownership	0.00	NA	NA
C3	Discovery Tool	0.00	NA	NA
C4	Production & Sharing	0.00	NA	NA
C5	Quality	0.33	NA	NA
C6	Ontology	0.00	NA	NA
C7	Archetypes	0.00	NA	NA
C8	Structural Components	0.00	NA	NA
C9	Lead Time	0.00	NA	NA
C10	Discoverability	0.00	NA	NA
C11	Addressability	0.00	NA	NA
C12	Trustworthiness	0.00	NA	NA
C13	Descriptiveness	0.00	NA	NA
C14	Interoperability	0.00	NA	NA
C15	Security	0.50	NA	NA

ID	Characteristic	Current Maturity Level	Target Maturity Level	Maturity Gap
C16	Accessibility	0.50	NA	NA
D1	Infrastructure & Platform	0.67	NA	NA
D2	Self-Serve	0.67	NA	NA
D3	Performance	1.00	NA	NA
D4	Ownership	0.33	NA	NA
D5	Platform Team	1.33	NA	NA
D6	Multiplane Platform	0.50	NA	NA
D7	Analytical API's	0.50	NA	NA
D8	Operational API's	0.33	NA	NA
E1	Security & Compliance	1.00	NA	NA
E2	Global Policies	1.00	NA	NA
E3	Federated Policies	0.33	NA	NA
E4	Monitoring	1.00	NA	NA
E5	Standardisation	0.50	NA	NA
E6	Computational Policies & Automation	0.50	NA	NA
E7	Governance Team	0.67	NA	NA
E8	Incident Management	1.00	NA	NA

Table E.2: Aggregated Responses Organisation II: Characteristics

ID	Characteristic	Current Maturity Level	Target Maturity Level	Maturity Gap
A1	Data-Oriented Strategy & Vision	2.33	3.00	0.67
A2	Culture, Mindset, & Values	1.00	3.00	2.00
A3	Value Realisation	1.00	3.00	2.00
A4	Curiosity & Ability	2.00	3.00	1.00
A5	Agile	1.33	3.00	1.67
A6	Executive Commitment	1.33	3.00	1.67
A7	Solid Engineering	2.00	3.00	1.00
A8	Change Management	1.00	3.00	2.00
A9	Value Adding Use Cases	1.67	3.00	1.33
A10	Roles	2.33	3.00	0.67
A11	Skills & Capabilities	1.33	3.00	1.67
A12	Incentivisation	0.33	3.00	2.67
A13	Training	1.33	3.00	1.67
B1	Definition	2.00	2.50	0.50
B2	Structure	2.33	3.00	0.67
B3	Decentralisation	1.67	3.00	1.33
B4	Ownership	1.33	3.00	1.67
B5	Autonomy	1.33	3.00	1.67
B6	Cross-Functional Teams	1.67	3.00	1.33
B7	Architecture	1.67	3.00	1.33
B8	Producers	1.33	3.00	1.67
B9	Consumers	1.33	3.00	1.67
C1	Definition	1.00	3.00	2.00
C2	Ownership	1.33	3.00	1.67
C3	Discovery Tool	1.00	3.00	2.00
C4	Production & Sharing	0.33	3.00	2.67
C5	Quality	1.67	3.00	1.33
C6	Ontology	0.67	3.00	2.33
C7	Archetypes	0.67	3.00	2.33
C8	Structural Components	0.33	2.50	2.17
C9	Lead Time	0.67	2.50	1.83
C10	Discoverability	1.00	3.00	2.00
C11	Addressability	1.00	3.00	2.00
C12	Trustworthiness	1.00	3.00	2.00

ID	Characteristic	Current Maturity Level	Target Maturity Level	Maturity Gap
C13	Descriptiveness	1.00	3.00	2.00
C14	Interoperability	1.33	3.00	1.67
C15	Security	1.67	3.00	1.33
C16	Accessibility	1.67	3.00	1.33
D1	Infrastructure & Platform	1.33	3.00	1.67
D2	Self-Serve	1.33	3.00	1.67
D3	Performance	1.33	3.00	1.67
D4	Ownership	2.00	3.00	1.00
D5	Platform Team	2.33	3.00	0.67
D6	Multiplane Platform	0.67	3.00	2.33
D7	Analytical API's	1.33	3.00	1.67
D8	Operational API's	1.67	3.00	1.33
E1	Security & Compliance	2.00	3.00	1.00
E2	Global Policies	1.33	3.00	1.67
E3	Federated Policies	1.33	3.00	1.67
E4	Monitoring	1.00	3.00	2.00
E5	Standardisation	1.33	3.00	1.67
E6	Computational Policies & Automation	0.67	3.00	2.33
E7	Governance Team	2.33	3.00	0.67
E8	Incident Management	1.67	3.00	1.33

Table E.3: Aggregated Responses Organisation III: Characteristics

ID	Characteristic	Current Maturity Level	Target Maturity Level	Maturity Gap
A1	Data-Oriented Strategy & Vision	2.50	NA	NA
A2	Culture, Mindset, & Values	0.50	NA	NA
A3	Value Realisation	1.00	NA	NA
A4	Curiosity & Ability	1.50	NA	NA
A5	Agile	1.00	NA	NA
A6	Executive Commitment	2.50	NA	NA
A7	Solid Engineering	1.00	NA	NA
A8	Change Management	1.00	NA	NA
A9	Value Adding Use Cases	2.00	NA	NA
A10	Roles	2.00	NA	NA
A11	Skills & Capabilities	0.00	NA	NA
A12	Incentivisation	0.00	NA	NA
A13	Training	0.00	NA	NA
B1	Definition	1.00	NA	NA
B2	Structure	2.00	NA	NA
B3	Decentralisation	1.00	NA	NA
B4	Ownership	1.00	NA	NA
B5	Autonomy	1.00	NA	NA
B6	Cross-Functional Teams	1.00	NA	NA
B7	Architecture	1.00	NA	NA
B8	Producers	1.00	NA	NA
B9	Consumers	1.50	NA	NA
C1	Definition	0.00	NA	NA
C2	Ownership	0.00	NA	NA
C3	Discovery Tool	0.00	NA	NA
C4	Production & Sharing	0.50	NA	NA
C5	Quality	0.50	NA	NA
C6	Ontology	1.00	NA	NA
C7	Archetypes	1.00	NA	NA
C8	Structural Components	0.00	NA	NA
C9	Lead Time	0.00	NA	NA
C10	Discoverability	0.50	NA	NA

ID	*Characteristic*	*Current Maturity Level*	*Target Maturity Level*	*Maturity Gap*
C11	Addressability	0.50	NA	NA
C12	Trustworthiness	0.00	NA	NA
C13	Descriptiveness	0.00	NA	NA
C14	Interoperability	0.00	NA	NA
C15	Security	0.50	NA	NA
C16	Accessibility	1.00	NA	NA
D1	Infrastructure & Platform	1.00	NA	NA
D2	Self-Serve	1.00	NA	NA
D3	Performance	1.00	NA	NA
D4	Ownership	0.00	NA	NA
D5	Platform Team	2.00	NA	NA
D6	Multiplane Platform	0.00	NA	NA
D7	Analytical API’s	0.00	NA	NA
D8	Operational API’s	3.00	NA	NA
E1	Security & Compliance	1.00	NA	NA
E2	Global Policies	1.00	NA	NA
E3	Federated Policies	0.00	NA	NA
E4	Monitoring	0.00	NA	NA
E5	Standardisation	0.00	NA	NA
E6	Computational Policies & Automation	0.00	NA	NA
E7	Governance Team	1.00	NA	NA
E8	Incident Management	0.50	NA	NA

F

Document Manual

Appendix F provides the maturity assessment document manual. The manual presents the different available sheets in the document `Data Mesh Maturity Assessment.xlsx` and explains each sheet's main purpose.

Manual

The overall structure of the maturity assessment model takes the form of seven sheets.

I. Overview	Provides an introduction and presents the overall structure of this document.
II. Assessment	Provides the assessment. This sheet guides you through all the characteristics by questions and definitions for the different maturity levels.
III. Responses	Provides the main statistics and an overview of the results.
IV. Responses PPT	Provides the main statistics and an overview of the results from *People*, *Process* and *Technology* perspectives.
V. Experiment	Enables the possibility to set different weights for the different dimensions and characteristics. Provides the updated statistics.
VI. Insights Radar	Provides insights in the assessment results by showing a radar chart. Filters are added which could be used to focus on specific dimensions.
VII. Insights Bar	Provides insights in the assessment results by showing a bar chart. Filters are added which could be used to focus on specific dimensions.

Page 2

I. Overview

Provides an introduction and presents the overall structure of this document.

Data Mesh Maturity Assessment

[Fill in Client Name]

Introduction:

This document provides a data mesh maturity assessment model.
The model is designed to enable the assessment how mature a data mesh implementation is within an organisation.
Five dimensions are considered to assess the data mesh maturity. These five dimensions are represented by, in total, 54 characteristics.

The dimensions are the following:

1. Data Foundation & Organisational Change (13)

2. Domain-Oriented Decentralised Data Ownership and Architecture (9)

3. Data as a Product (16)

4. Self-Serve Data Infrastructure as a Platform (8)

5. Federated Computational Governance (8)

Structure:

The overall structure of this document takes the form of seven sheets. You could use the buttons.

I. Overview
Provides an introduction and presents the overall structure of this document.

II. Assessment
Provides the assessment. This sheet guides you through all the characteristics by questions and definitions for the different maturity levels.

III. Responses
Provides the main statistics and an overview of the results.

IV. Responses PPT
Provides the main statistics and an overview of the results from *People*, *Process* and *Technology* perspectives.

V. Experiment
Enables the possibility to set different weights for the different dimensions and characteristics. Provides the updated statistics.

VI. Insights Radar
Provides insights in the assessment results by showing a radar chart. Filters are added which could be used to focus on specific dimensions.

VII. Insights Bar
Provides insights in the assessment results by showing a bar chart. Filters are added which could be used to focus on specific dimensions.

Page 3

II. Assessment

Provides the assessment. This sheet guides you through all the characteristics by questions and definitions for the different maturity levels.

Data Mesh Maturity Assessment

Instructions:

Please rate the following 54 questions on the current and target level of maturity that you find most fitting for your organisation.

Each statement should be rated on a scale from level 0 "non-initiated" to level 3 "achieved", or as "unknown".

The descriptions guide you in making your choice.

The maturity levels are defined as:	**Level 0** "Non-initiated"	**Level 1** "Conceptual"	**Level 2** "Defined"	**Level 3** "Achieved"

If you have any comments, please enter them in the comments section.

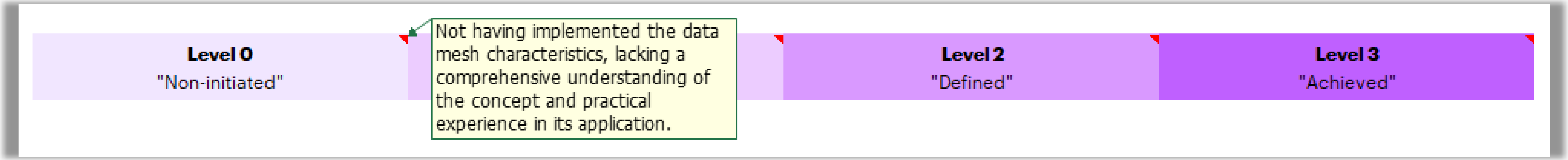
A	Data Foundation & Organisational Change						
	Organisational data fundamentals and the needed changes introduced to people's roles, responsibilities, motivations, and collective interactions in an organisation.						
A1	**Data-Oriented Strategy & Vision**	**0**	**1**	**2**	**3**	**Current State**	**Target State**	**Comments**
	Is your organisation adopting a data-oriented strategy, based on a vision to use data?	No effective data-oriented strategy and vision considered. Lack of understanding on how data will drive business growth. No roadmaps defined for the development of data initiatives.	Initial data-oriented strategy and vision created. Ad hoc data & analytics initiatives developed. Often not considered to be aligned with initially created data strategy. Target states not determined.	Comprehensive data-oriented strategy and vision in place. Data & analytics and machine learning initiatives are developed where organisational data strategy is backing the decisions. Data initiatives serve the overall business interest. Roadmaps created to achieve target states.	Effective data-oriented strategy and vision adopted. Data is treated as strategic asset, which directly drives business growth. Strategy drives gaining competitive advantage by using data & analytics and machine learning as differentiators. Current state assessments performed to create roadmaps to achieve target states for all functions.			
A2	**Culture, Mindset & Values**	**0**	**1**	**2**	**3**	**Current State**	**Target State**	**Comments**
		Data as an asset to be collected and	Cross-domain data sharing initiatives set up. Only few people understand how data	Real data product-thinking mindset. Data is	Cultural shift from data protection towards data sharing mindset accomplished. Data as a			

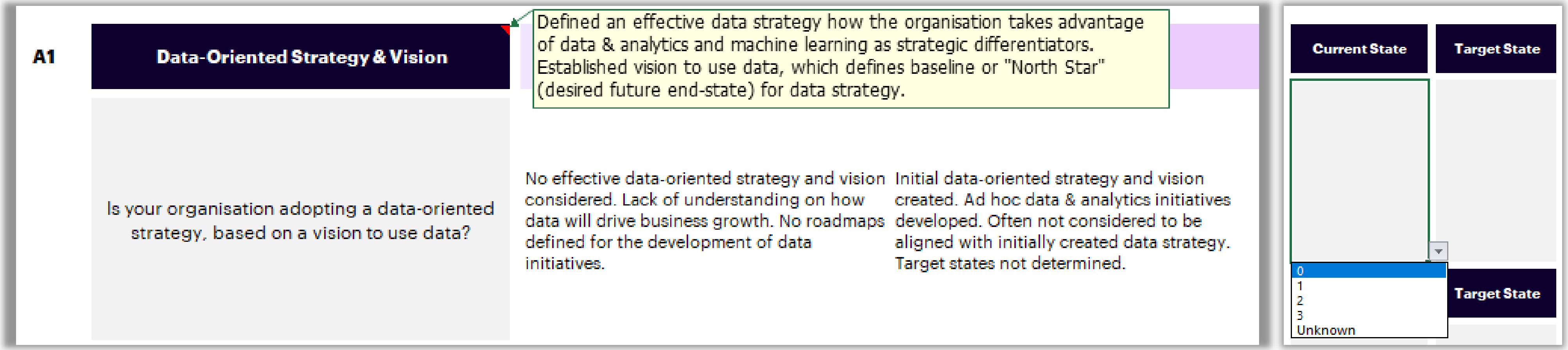
Page 4

Definitions are added by comment boxes for all maturity levels and characteristics.
Drop-down functions are added to select the maturity levels 0, 1, 2, 3 or to select Unknown.

II. Assessment

Provides the assessment. This sheet guides you through all the characteristics by questions and definitions for the different maturity levels.





Page 5

After conducting the self-assessment, the results will automatically be calculated for the relevant metrics. Definitions for all metrics are provided by comment boxes.

III. Responses

Provides the main statistics and an overview of the results.

Statistics:

ID	Dimension	Current Maturity Level	Unknown Rate	Response Rate	Target Maturity Level	Unknown Rate	Response Rate	Maturity Gap
A-E	All Dimensions			0%			0%	
A	Data Foundation & Organisational Change			0%			0%	
B	Domain-Oriented Decentralised Data Ownership & Architecture			0%			0%	
C	Data as a Product			0%			0%	
D	Self-Serve Data Infrastructure as a Platform			0%			0%	
E	Federated Computational Governance			0%			0%	

Statistics:

ID	Dimension	Current Maturity Level	Unknown Rate	Response Rate	Target Maturity Level	Unknown Rate	Response Rate	Maturity Gap
A-E	All Dimensions	0.811	2%	100%	2.412	6%	100%	1.600
A	Data Foundation & Organisational Change	1.000	8%	100%	2.583	8%	100%	1.583
B	Domain-Oriented Decentralised Data Ownership & Architecture	0.889	0%	100%	2.571	22%	100%	1.683
C	Data as a Product	0.688	0%	100%	2.125	0%	100%	1.438
D	Self-Serve Data Infrastructure as a Platform	0.750	0%	100%	2.375	0%	100%	1.625
E	Federated Computational Governance	0.750	0%	100%	2.625	0%	100%	1.875

Unknown Rate		
2%		
8%	100%	2.583
0%	100%	2.571

The number of questions where the current maturity level is answered by "Unknown" relative to the total number of questions where the current maturity level is provided. [Percentage]

Page 6

An overview of all the provided answers will be presented. In addition, maturity gaps will be calculated for each characteristic.

III. Responses

Provides the main statistics and an overview of the results.

Responses:

ID	Dimension	Characteristic	Current	Target	Maturity Gap	Comments
A1	Data Foundation & Organisational Change	Data-Oriented Strategy & Vision	0	1	1	
A2	Data Foundation & Organisational Change	Culture, Mindset & Values	1	2	1	
A3	Data Foundation & Organisational Change	Value Realisation	1	2	1	
A4	Data Foundation & Organisational Change	Curiosity & Ability	1	3	2	
A5	Data Foundation & Organisational Change	Agile	0	2	2	
A6	Data Foundation & Organisational Change	Executive Commitment	1	3	2	
A7	Data Foundation & Organisational Change	Solid Engineering	1	3	2	
A8	Data Foundation & Organisational Change	Change Management	Unknown	Unknown		
A9	Data Foundation & Organisational Change	Value Adding Use Cases	1	3	2	
A10	Data Foundation & Organisational Change	Roles	2	3	1	
A11	Data Foundation & Organisational Change	Skills & Capabilities	2	3	1	
A12	Data Foundation & Organisational Change	Incentivisation	1	3	2	
A13	Data Foundation & Organisational Change	Training	1	3	2	
B1	Domain-Oriented Decentralised Data Ownership & Architecture	Definition	0	3	3	
B2	Domain-Oriented Decentralised Data Ownership & Architecture	Structure	2	3	1	
B3	Domain-Oriented Decentralised Data Ownership & Architecture	Decentralisation	2	3	1	
B4	Domain-Oriented Decentralised Data Ownership & Architecture	Ownership	1	2	1	
B5	Domain-Oriented Decentralised Data Ownership & Architecture	Autonomy	1	Unknown		
B6	Domain-Oriented Decentralised Data Ownership & Architecture	Cross-Functional Teams	0	Unknown		
B7	Domain-Oriented Decentralised Data Ownership & Architecture	Architecture	0	3	3	
B8	Domain-Oriented Decentralised Data Ownership & Architecture	Producers	1	2	1	
B9	Domain-Oriented Decentralised Data Ownership & Architecture	Consumers	1	2	1	
C1	Data as a Product	Definition	0	3	3	
C2	Data as a Product	Ownership	0	3	3	
C3	Data as a Product	Discovery Tool	1	3	2	
C4	Data as a Product	Production & Sharing	1	2	1	
C5	Data as a Product	Quality	2	2	0	
C6	Data as a Product	Ontology	2	2	0	
C7	Data as a Product	Archetypes	1	2	1	
C8	Data as a Product	Structural Components	0	3	3	
C9	Data as a Product	Lead Time	0	3	3	
C10	Data as a Product	Discoverability	1	3	2	

Page 7

After conducting the self-assessment, the results will automatically be calculated for the relevant metrics approached from the different perspectives.

IV. Responses PPT

Provides the main statistics and an overview of the results from *People*, *Process* and *Technology* perspectives.

Statistics:
People , *Process* , *Technology* Perspectives

Nr.	Dimension	Current Maturity Level	Unknown Rate	Response Rate	Target Maturity Level	Unknown Rate	Response Rate	Maturity Gap
I	People			0%			0%	
II	Process			0%			0%	
III	Technology			0%			0%	

Statistics:
People , *Process* , *Technology* Perspectives

Nr.	Dimension	Current Maturity Level	Unknown Rate	Response Rate	Target Maturity Level	Unknown Rate	Response Rate	Maturity Gap
I	People	0.879	3%	100%	2.387	9%	100%	1.508
II	Process	0.766	2%	100%	2.370	4%	100%	1.604
III	Technology	0.714	5%	100%	2.476	5%	100%	1.762

Page 8

People, Process, and Technology labels are provided for all the characteristics.

IV. Responses PPT

Provides the main statistics and an overview of the results from *People*, *Process* and *Technology* perspectives.

Responses:
Characteristics classified along *People*, *Process* and *Technology* perspectives.

ID	Dimension	Characteristic	People	Process	Technology	Current	Target	Maturity Gap
A1	Data Foundation & Organisational Change	Data-Oriented Strategy & Vision	People	Process	Technology			
A2	Data Foundation & Organisational Change	Culture, Mindset & Values	People					
A3	Data Foundation & Organisational Change	Value Realisation		Process	Technology			
A4	Data Foundation & Organisational Change	Curiosity & Ability	People		Technology			
A5	Data Foundation & Organisational Change	Agile	People	Process	Technology			
A6	Data Foundation & Organisational Change	Executive Commitment	People	Process				
A7	Data Foundation & Organisational Change	Solid Engineering			Technology			
A8	Data Foundation & Organisational Change	Change Management	People	Process	Technology			
A9	Data Foundation & Organisational Change	Value Adding Use Cases		Process	Technology			
A10	Data Foundation & Organisational Change	Roles	People					
A11	Data Foundation & Organisational Change	Skills & Capabilities	People					
A12	Data Foundation & Organisational Change	Incentivisation	People	Process				
A13	Data Foundation & Organisational Change	Training	People	Process	Technology			
B1	Domain-Oriented Decentralised Data Ownership & Architecture	Definition		Process				
B2	Domain-Oriented Decentralised Data Ownership & Architecture	Structure	People	Process				
B3	Domain-Oriented Decentralised Data Ownership & Architecture	Decentralisation	People	Process				
B4	Domain-Oriented Decentralised Data Ownership & Architecture	Ownership	People	Process				
B5	Domain-Oriented Decentralised Data Ownership & Architecture	Autonomy	People	Process				
B6	Domain-Oriented Decentralised Data Ownership & Architecture	Cross-Functional Teams	People					
B7	Domain-Oriented Decentralised Data Ownership & Architecture	Architecture	People	Process	Technology			
B8	Domain-Oriented Decentralised Data Ownership & Architecture	Producers	People	Process				
B9	Domain-Oriented Decentralised Data Ownership & Architecture	Consumers	People	Process				
C1	Data as a Product	Definition		Process				
C2	Data as a Product	Ownership	People	Process				
C3	Data as a Product	Discovery Tool	People	Process	Technology			
C4	Data as a Product	Production & Sharing	People	Process				
C5	Data as a Product	Quality		Process				
C6	Data as a Product	Ontology	People	Process				
C7	Data as a Product	Archetypes		Process				
C8	Data as a Product	Structural Components		Process				
C9	Data as a Product	Lead Time		Process				
C10	Data as a Product	Discoverability	People	Process				
C11	Data as a Product	Addressability		Process				
C12	Data as a Product	Trustworthiness	People	Process				
C13	Data as a Product	Descriptiveness		Process				
C14	Data as a Product	Interoperability	People	Process				

Page 9

Assign *critical* and *non-critical* dimensions by setting different weights.

V. Experiment

Enables the possibility to set different weights for the different dimensions and characteristics. Provides the updated statistics.

Statistics:

Define different importance levels (weights) of the different **dimensions** below.
Exclude, Low, Medium, High

ID	Dimension	Current Maturity Level	Target Maturity Level	Maturity Gap
A-E	All Dimensions			

ID	Dimension	Importance
A	Data Foundation & Organisational Change	Medium
B	Domain-Oriented Decentralised Data Ownership & Architecture	Medium
C	Data as a Product	Medium
D	Self-Serve Data Infrastructure as a Platform	Medium
E	Federated Computational Governance	Medium

Level	Weight
Exclude	0
Low	1
Medium	2
High	3

Page 10

Assign *critical* and *non-critical* characteristics by setting different weights.

V. Experiment

Enables the possibility to set different weights for the different dimensions and characteristics. Provides the updated statistics.

Define different importance levels (weights) of the different **characteristics** below.

Exclude, Low, Medium, High

ID	Dimension	Current Maturity Level	Target Maturity Level	Maturity Gap
A-E	All Dimensions			
A	Data Foundation & Organisational Change			
B	Domain-Oriented Decentralised Data Ownership & Architecture			
C	Data as a Product			
D	Self-Serve Data Infrastructure as a Platform			
E	Federated Computational Governance			

Level	Weight
Exclude	0
Low	1
Medium	2
High	3

ID	Dimension	Characteristic	Importance
A1	Data Foundation & Organisational Change	Data-Oriented Strategy & Vision	Medium
A2	Data Foundation & Organisational Change	Culture, Mindset & Values	Medium
A3	Data Foundation & Organisational Change	Value Realisation	Medium
A4	Data Foundation & Organisational Change	Curiosity & Ability	Medium
A5	Data Foundation & Organisational Change	Agile	Medium
A6	Data Foundation & Organisational Change	Executive Commitment	Medium
A7	Data Foundation & Organisational Change	Solid Engineering	Medium
A8	Data Foundation & Organisational Change	Change Management	Medium
A9	Data Foundation & Organisational Change	Value Adding Use Cases	Medium
A10	Data Foundation & Organisational Change	Roles	Medium
A11	Data Foundation & Organisational Change	Skills & Capabilities	Medium
A12	Data Foundation & Organisational Change	Incentivisation	Medium
A13	Data Foundation & Organisational Change	Training	Medium
B1	Domain-Oriented Decentralised Data Ownership & Architecture	Definition	Medium
B2	Domain-Oriented Decentralised Data Ownership & Architecture	Structure	Medium
B3	Domain-Oriented Decentralised Data Ownership & Architecture	Decentralisation	Medium
B4	Domain-Oriented Decentralised Data Ownership & Architecture	Ownership	Medium
B5	Domain-Oriented Decentralised Data Ownership & Architecture	Autonomy	Medium
B6	Domain-Oriented Decentralised Data Ownership & Architecture	Cross-Functional Teams	Medium
B7	Domain-Oriented Decentralised Data Ownership & Architecture	Architecture	Medium
B8	Domain-Oriented Decentralised Data Ownership & Architecture	Producers	Medium
B9	Domain-Oriented Decentralised Data Ownership & Architecture	Consumers	Medium
C1	Data as a Product	Definition	Medium
C2	Data as a Product	Ownership	Medium
C3	Data as a Product	Discovery Tool	Medium
C4	Data as a Product	Production & Sharing	Medium
C5	Data as a Product	Quality	Medium
C6	Data as a Product	Ontology	Medium

Page 11

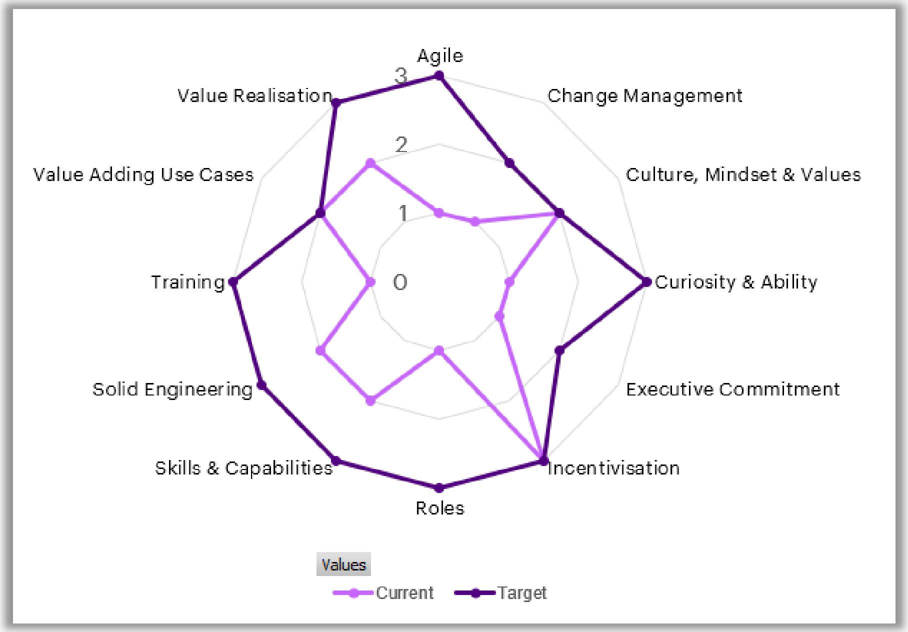
Visual representation of the results by using a radar chart.

VI. Insights Radar

Provides insights in the assessment results by showing a radar chart. Filters are added which could be used to focus on specific dimensions.

Dimension	Data Foundation & Organisational Change
Current	(Multiple Items)
Target	(Multiple Items)

Characteristic	Current	Target
Agile	1	3
Change Management	1	2
Culture, Mindset & Values	2	2
Curiosity & Ability	1	3
Executive Commitment	1	2
Incentivisation	3	3
Roles	1	3
Skills & Capabilities	2	3
Solid Engineering	2	3
Training	1	3
Value Adding Use Cases	2	2
Value Realisation	2	3
Grand Total	3	3



Page 12

Visual representation of the results by using a bar chart.

VII. Insights Bar

Provides insights in the assessment results by showing a bar chart. Filters are added which could be used to focus on specific dimensions.