



**HAL**  
open science

# Data Management Maturity Model-Process Dimensions and Capabilities to Leverage Data-Driven Organizations Towards Industry 5.0

Lara Pörtner, Andreas Riel, Benedikt Schmidt, Marcel Leclaire, Robert Möske

## ► To cite this version:

Lara Pörtner, Andreas Riel, Benedikt Schmidt, Marcel Leclaire, Robert Möske. Data Management Maturity Model-Process Dimensions and Capabilities to Leverage Data-Driven Organizations Towards Industry 5.0. Applied System Innovation, 2025, 8 (2), pp.41. <10.3390/asi8020041>. <hal-05001461>

**HAL Id: hal-05001461**

**<https://hal.science/hal-05001461v1>**

Submitted on 21 Mar 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

## Article

# Data Management Maturity Model—Process Dimensions and Capabilities to Leverage Data-Driven Organizations Towards Industry 5.0

Lara Pörtner <sup>1</sup>, Andreas Riel <sup>1,\*</sup> , Benedikt Schmidt <sup>2</sup>, Marcel Leclaire <sup>2</sup> and Robert Möske <sup>2</sup>

<sup>1</sup> Grenoble Institute of Engineering and Management, Grenoble Alpes University, CNRS, Grenoble INP, G-SCOP, 46 Avenue Félix Viallet, 38000 Grenoble, France; lara.portner@grenoble-inp.fr

<sup>2</sup> CAMELOT Management Consultants AG, Gabrielenstraße 9, 80636 Munich, Germany; bene.sch.217@gmail.com (B.S.); marcel.leclaire@gmx.de (M.L.); moeske.robert@gmail.com (R.M.)

\* Correspondence: andreas.riel@grenoble-inp.fr

**Abstract:** Data-driven organizations aim to control business decisions based on data. However, despite significant investments in digitalization, studies show that many organizations continue to face challenges in fully realizing the benefits of data. Existing maturity models for digital transformation, data management, and data-driven organizations lack a comprehensive, industry-agnostic, and practically validated approach to addressing industry challenges. This work introduces a refined data management maturity model developed using De Bruin’s maturity model assessment methodology. The model aims to incorporate all key elements of a data-driven organization, emphasizing the interdependencies required to evaluate maturity levels and provide targeted recommendations for addressing data-related challenges during the transition to a data-driven organization. An initial validation with 31 industry experts confirmed the model’s feasibility and practical applicability. As the next step, we plan to validate the model further by deploying the full questionnaire and deriving the maturity of each process dimension, along with its weighting, through assessments with industry partners from various sectors, including automotive, aviation, consumer goods/manufacturing, pharma, and media. Preliminary findings also underscored the importance of a deeper focus on the organization dimension, particularly in the context of Industry 5.0. Future research will refine the model through iterative development phases to address this critical area.

**Keywords:** maturity model; capability assessment; data-driven organization



Academic Editor: Mario Di Nardo

Received: 28 December 2024

Revised: 8 February 2025

Accepted: 19 February 2025

Published: 21 March 2025

**Citation:** Pörtner, L.; Riel, A.; Schmidt, B.; Leclaire, M.; Möske, R.

Data Management Maturity Model—Process Dimensions and Capabilities to Leverage Data-Driven Organizations Towards Industry 5.0. *Appl. Syst. Innov.* **2025**, *8*, 41. <https://doi.org/10.3390/asi8020041>

**Copyright:** © 2025 by the authors. Published by MDPI on behalf of the International Institute of Knowledge Innovation and Invention. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The technology portfolio of Industry 4.0 includes innovations such as the Internet of Things (IoT), cloud computing, artificial intelligence, and machine learning. In all of these technologies, data play an indispensable role. Data provide the essential foundation on which many organizations increasingly base their decisions and processes, bringing them closer to becoming data-driven organizations [1]. However, only one in four companies in Germany currently benefits from data [2]. Since 2022, the industry has progressed further and is now discussing “Industry 5.0”, where the focus is on “digitalization with a purpose”. According to Paul et al. [3], Industry 5.0 can be viewed as an extension of Industry 4.0, with a stronger emphasis on human–machine interaction and collaboration. Industry 5.0, therefore, not only takes a more customer-centric approach but also adopts a broader human-centered view [4]. Unlike Industry 4.0, which is technology-driven, Industry 5.0 is

value-driven [5]. This shift underscores the need for organizations to harness the value of their data to remain competitive by focusing on relevant dimensions of the organization and its data. For digitalization to advance and for companies to extract value from data, organizations need a comprehensive digital strategy and roadmap to reach a mature level.

Maturity models are often used to assess the status quo of various capabilities within data-driven organizations. Their application serves as a foundation for deriving roadmaps and recommendations for action [6]. Existing maturity assessments related to data-driven organizations attempt to determine maturity levels in relation to data. These assessments aim to analyze an organization's current (as-is) situation by identifying strengths and weaknesses using maturity models. While the existing literature largely discusses the association between processes, capabilities, and data-driven value creation, questions remain about which capabilities should be included in a data management maturity model to foster data-driven organizations. It is also unclear what potential relationships exist between individual capabilities and how these relationships influence the overall maturity level of the organization and the roadmap for improvement.

Additionally, it is important to assess whether companies are prepared for the next stage, "Industry 5.0", which focuses more on the human aspects of organizations. In the context of data-driven organizations, generating value from data by focusing on societal components becomes increasingly relevant. However, value and technology must work in tandem within industry concepts [5]. Given the complexity of large organizations, broadening the scope within maturity models rather than focusing on a single capability dimension could lead to better outcomes in developing roadmaps toward a fully data-driven organization.

This raises the following questions: What do data-driven organizations comprise? How can the value of data be best exploited? And to what extent has the "people" dimension gained acceptance within organizations?

To address these questions, a data management maturity model used in practice was taken as a reference model to compare capabilities with existing studies from the research. Based on a literature review, an industry survey, and expert knowledge, a refined maturity model is herein proposed. This model accounts for the necessary correlations when assessing maturity, allowing for more precise recommendations to manage data challenges.

This work attempts to answer the following research question:

RQ: Which capabilities and process dimensions should be part of a data management maturity model that can be used in practice as a basis for interviews to determine the status quo in the transition to a data-driven organization?

To achieve the objective of this study, the paper proceeds as follows: Section 2 defines key terms related to the background and provides an overview of the methodology. It also examines the current research status of maturity models and identifies the research gap. Section 3 outlines the development phases of the maturity assessment in detail, followed by a discussion of the findings and limitations. Section 4 also offers an outlook on future research plans.

## 2. Background and Methodology

### 2.1. Data-Driven Organization

The exploitation of insights derived from data in all strategic business decisions characterize data-driven organizations [1]. Moreover, data science [1] plays an important role in data-driven decision making to extract valuable insights from data and improve the business [2].

To become data-driven, the data and technology capabilities of an organization are not the only significant indicators. Also, organizational capabilities in terms of a data culture are important in the journey to becoming a data-driven organization [3].

Hupperz et al. [4] introduced a joint map with different necessary key elements that form a data-driven organization (Table 1). This includes key elements: digital transformation, data science, data-driven business model, data-driven innovation, and data analytics. The key element of digital transformation looks at the organization itself with its strategy, culture, and the entire business ecosystem. Data science focuses on how IT resources create the necessary value from data to drive competitiveness. In this context, the key element of the data-driven business model covers new business models, and the way in which values are created. Data analytics is aimed at the analytical skills of being able to analyze and evaluate data that are introduced due to data-driven innovation possibilities such as big data and business insights.

**Table 1.** Key elements of data-driven organizations (Hupperz et al. [4]).

Key Elements		
Data-Driven Organization	Digital Transformation (K1)	Strategy (C11) Data-driven culture (C12) Business ecosystem (C13)
	Data Science (K2)	IT Professionals (C21) Add value (C22) Competitive advantage (C23)
	Data-Driven Business Model (K3)	Demand (C31) BM development (C32) Value creation (C33)
	Data-Driven Innovation (K4)	Big data (C41) R&D (C42) Business insights (C43)
	Data Analytics (K5)	Descriptive (C51) Predictive (C52) Prescriptive (C53)

### 2.2. Maturity Model

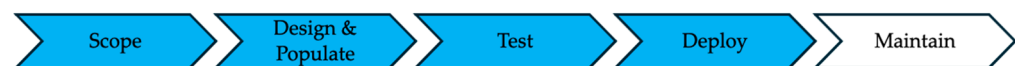
Maturity models (MM) aim to assist organizations by providing guidance with their sequence of discrete maturity levels for a class of processes [5]. With their structural approaches, they offer a roadmap for the organization’s improvement.

The Capability Maturity Model Integration (CMMI) and Software Process Improvement and Capability dEtermination (SPICE) are a set of standards, i.e., ISO/IEC 330xx, known as fundamental patterns in the assessment of process capabilities [5–7]. CMMI has been a major standard for software and systems development for more than 20 years. They can be used descriptively to assess the current process capability level. These models have been customized to different domains, including the automotive industry and government [8].

The data management maturity model (DMM) consists six functional domains (data management strategy, data quality, data operations, platform and architectures, data governance, and supporting processes) and was developed using the principles and structure of CMMI [6]. The maturity model was introduced in Pörtner et al. 2022 [9]. It provides an overview of the maturity level of data management in an organization [10,11] with a strong master-data focus. Twenty-six capabilities are defined under six process groups to provide maturity levels with a relevance evaluation for creating an extensive roadmap for continuous improvement in a structured and repeatable way. Master data in that context are a specific set of data that describes the key business entities on which an organization’s activities are based, such as business partners, products, or employees [12]. Moreover, objects of master data can be seen as official and basic business objects in companies that are used in value-adding processes [13].

### 2.3. Methodology

To answer the research question, this work follows De Bruin's six phases for developing a maturity assessment [14]. The approach typically begins with a systematic literature review to establish a clearer understanding of how to structure the maturity model. In this work, the development process starts with an existing holistic data management maturity model used as a reference model. The results and implications of this iteration will inform future research aimed at developing a maturity model with a stronger focus on people and organizations and aligned with the principles of Industry 5.0, to guide the transition to data-driven organizations. Figure 1 illustrates the phases that were followed. The subsequent sections provide a detailed description of the development model, presenting results for each phase.



**Figure 1.** Maturity Assessment Development Phases adapted on De Bruin [14].

The aim of the scope phase was to define the model. Using an existing reference model, the focus was placed on an industry-specific and practice-oriented model that addresses the specific challenges and requirements of the respective domain.

The structure of the model was developed in the design phase. The structure and maturity levels were already defined by the existing reference model and were only adapted based on findings from the literature review and gaps to existing research.

In this work, the populate phase goes hand in hand with the design phase. This is where the relevant content of the model was finally defined, which captures the maturity level of the domain.

The model was then validated in the test phase using results from an industry survey. In addition, experts from the field of data management validated possible correlations between the dimensions.

In the subsequent deploy phase, interviews and pilot tests were conducted for relevance, validity, and comprehensibility. As part of the short assessment, which contains a small number of questions per dimension from the questionnaire, further input was obtained. The participants in the industry survey and the short assessment were different. Nevertheless, care was taken to ensure that different industry sectors were represented. For the participants in the short assessment in particular, care was also taken to ensure that they had in-depth expertise in the field of data and analytics and that their answers were representative of the respective industry.

For long-term use, the model will be tested as part of the maintain phase with the comprehensively compiled questionnaire. The model will thus be regularly reviewed and adapted to new requirements to maintain its relevance and ensure comparability. Continuous further development is part of future research, as is the creation of a guide for applying the model in practice.

## 3. Data Management Maturity Model

### 3.1. Phase 1—Scope

This phase defines the focus of the model, building on an existing reference model and setting a domain-specific and practical approach.

#### 3.1.1. Model Context and Objectives

The data management maturity model [9], which has been applied in practice for several years, serves as a reference model in the development process of the maturity

assessment in this work. This model was further refined by incorporating research findings, results from industry surveys, and expert insights. The proposed version of the model highlights challenges and deep-dive topics that need to be addressed in future work, using explicit maturity models to support organizations on their journey toward becoming data-driven.

The reference model is domain-specific and aims to provide a holistic overview, focusing specifically on master-data management. Grounded in a 2012 study on strategic master-data management [15], the model functions as an assessment tool to help organizations understand their current (as-is) situation and identify potential areas for improvement.

In scope of this work is the refinement of the maturity map, which serves as part of the assessment to record the as-is situation of an organization by addressing different capabilities with the help of questions and assigning them to a maturity level. It does not describe how the assessment approach and its applicability can look in practice.

### 3.1.2. Definition of Model Focus

The refined maturity model provides an overview of the maturity level of data management within an organization. It also facilitates benchmarking against other companies and serves as a foundation for developing a data management roadmap as a next step. The reference model has already been applied across various industry sectors and has a strong focus on master data that need to be shifted to data in general.

The model categorizes maturity into six levels. Additionally, a relevance status is assigned to each maturity level to guide strategy and roadmap development. This relevance status is based on a scale from 1 to 4, reflecting customer preferences and priorities regarding future implementations. The process dimension is organized into six pillars of (master) data management: strategy, processes, organization, master data, IT systems, and operation and support, encompassing a total of 26 capabilities in total that more precisely elaborate the process dimensions (cf. Figure 2).

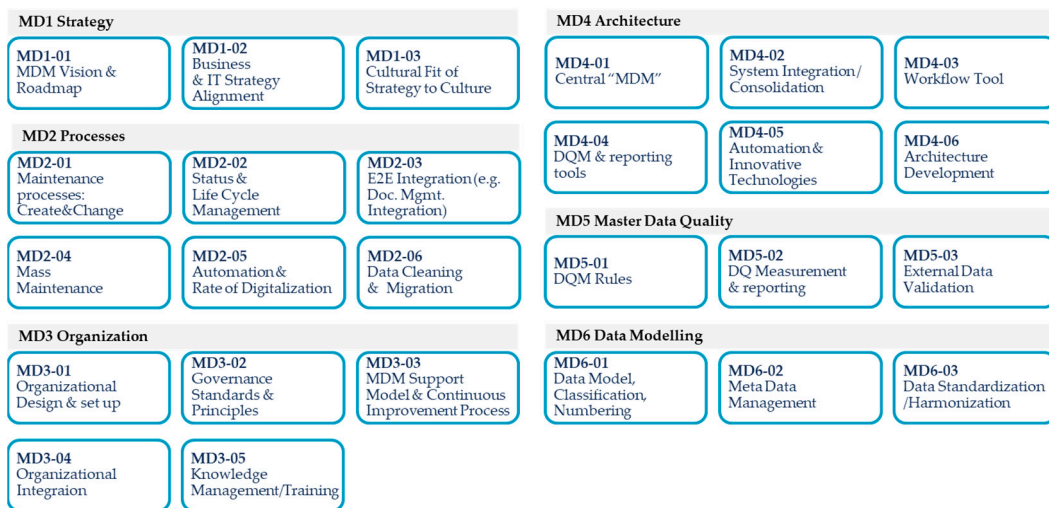


Figure 2. Reference Maturity Map [9].

### 3.1.3. Definition of Stakeholders

Due to the domain-specific focus of the reference model in the area of data management, special attention was given to ensure that the industry experts and other stakeholders involved in the model development process came from the field of data and analytics and had expertise in data management. Additionally, when selecting participants for the study, care was taken to ensure representation across different industry sectors—automotive, aviation, consumer goods/manufacturing, pharma, and media—by including participants

with specific expertise in these areas. The insights and knowledge of these experts were used to refine the model design and validate its components. Consequently, the 31 survey participants represent an initial cross-section of expert opinions. Based on the expertise of the experts, the sample size is assumed to be a representative sample.

The refined model is targeted at organizations across all industry sectors, with a particular focus on data project teams that include participants from both operational and management levels. The assessors for the evaluation were expected to be consultants or educators with at least basic knowledge in data management.

#### 3.1.4. Initial Literature Review

To capture and evaluate the current state of research, a literature review was conducted to assess the status of existing maturity models related to digital transformation, data management, and data-driven organizations. The aim was to provide an initial overview of which models exist in research and practice and to identify any gaps by also comparing the reference model.

With the significant increase in data in the 21st century [16] and the shift toward digital transformation, organizations now require structured approaches to navigate this transition. Maturity models (MMs) are commonly used to reflect an organization's initial status in this process [13]. By 2023, substantial research had been conducted in digital transformation. However, "most studies focus primarily on the overall process rather than on how companies develop a strategy to effectively utilize their data for Digital Transformation", and "maturity models have been largely neglected" [13].

Since MMs are still in the early stages of development, and only 27 percent of models have been published in academic journals, Gökalp and Martinez [8] proposed the digital transformation capability maturity model (DX-CMM) in 2021 to address this gap. They evaluated 18 MMs based on criteria such as "complete, clear, unambiguous, objective, impartial, consistent, repeatable, comparable, and representative" and developed DX-CMM as a holistic model applicable across all industry sectors. The model encompasses two dimensions—processes and capabilities—that were already applied in a case study. The model identifies an organization's maturity level in digital transformation and highlights opportunities for improvement.

Since the work of Gökalp and Martinez in 2021, the field of digital transformation has continued to evolve. Spruit and Pietzka [11] introduced the MD3M maturity model, which focuses on master data. Their work provides an overview of existing master-data management models and developed a model that can assess an organization's maturity in master-data management. The model covers five key topics: data model, data quality, usage and ownership, data protection, and maintenance. However, this model is not classified as an MM for strategic guidance in digital transformation, and it has not yet been widely adopted in practice. This reinforces the need for a practical, domain-specific model that supports assessments in digital transformation.

Another work by Gökalp et al. [6] introduced a multidisciplinary assessment approach for data-driven organizations, called the data-drivenness process capability determination model (DDPCDM). This model was validated for its applicability and usability through a multiple-case study involving two organizations. However, it has not yet been tested in small- and medium-sized enterprises (SMEs), and its practical applicability across all industry sectors remains to be demonstrated.

The need for a data management maturity model has also arisen in the automotive industry, which led to the presentation of the pilot draft of the data management SPICE assessment by the intacs group in 2022 [17]. Data management SPICE defines the fundamental business processes of data management by establishing a process assessment model (PAM) for the maturity of all types of data and data management processes within organizations.

In the work of Pörtner et al. [13], a comparison was made between the data management strategy maturity model, which is used as reference model in this work, and the digital transformation capability maturity model, based on the criteria outlined in ISO 33020.

### 3.1.5. Definition of Research Objective and Question

Based on the literature review and comparison analysis described in Section 3.1.4, the following hypotheses and key requirements arise for developing a domain-specific refined maturity model. Existing research highlights a lack of relevant empirical studies involving real-world use cases and models that have been validated for practical application across all industry sectors. Therefore, a holistic domain-specific maturity model must be created that addresses challenges of all industry sectors and that is validated through initial iterations. The field of digital transformation is highly dynamic, and the increasing volume of data, along with data management challenges, can have a significant impact. As such, a maturity model needs to be agile and adaptable to evolving market conditions and changing organizational needs. Moreover, a clear questionnaire behind each capability should be provided to understand better the link between each capability and maturity level.

The reference model demonstrates that expert knowledge is essential for conducting the data management strategy assessment and accurately evaluating an organization's maturity level. Maturity levels should be clearly defined to enable less experienced assessors to apply the assessment, ensuring greater objectivity and consistency.

While the reference model provides a comprehensive overview of a company's data management maturity across various industries, it currently focuses heavily on master data. The scope of the maturity model should not be limited to a single data type (master data) but should be broadened to assess data management as a whole. Therefore, in this work, capabilities need to be adapted to the topic of data management in general.

### 3.2. Phases 2 and 3—Design and Populate

During the design phase, the structure of the model was developed based on the scope and boundaries defined in the initial scope phase. Since an existing maturity model was used as a reference, and the initial design was revised; there are clear overlaps between the design and populate phases, which are described together below. The goal of these phases was to update the initial maturity map based on research and expert input, addressing challenges related to the transition to a data-driven organization.

Additionally, the setup of the model, including the definition of maturity levels and their core components, needs to be established. Clear documentation for practical application should be prepared during the maintain phase to ensure the validation of the full assessment.

#### 3.2.1. Systematic Literature Review

An initial literature review was conducted to gain a preliminary understanding of maturity models in the context of digital transformation (Section 3.1.4). To delve deeper into maturity models within this specific research area over recent years, including their levels of detail, weighting of capabilities, and the capabilities used in existing studies, a systematic literature review was performed using the IEEE and Science Public databases. The search utilized the following keywords: (“data-driven” OR “data science” OR “data literacy”) AND (“maturity model” OR “capability model” OR “assessment model”). These keywords were adapted from Gökalp et al. [6], who had previously conducted a literature review on this topic. The search was limited to studies published within the last six years (January 2018 to January 2024).

As a result of this research, we retrieved 208 studies. Additionally, we included nine relevant studies from Gökalp and their work on maturity models, bringing the total to

217 studies. After removing duplicates, 214 unique studies were available for the initial review. The evaluation was based on the following criteria: (1) publication in a conference paper, book chapter, or journal; (2) publication in English; (3) proposal of a capability or maturity model for data-driven organizations; (4) proposal of specific capabilities; and (5) description of an implementation strategy or journey roadmap. After the first review process, 32 studies were identified as relevant for this research.

In the second step, these 32 studies were further analyzed based on the following criteria: (1) maturity measurement, (2) use of capabilities to describe data-driven organizations, (3) concept of analyzing both the as-is and to-be status, (4) development of a roadmap with recommendations for action, (5) use of an assessment tool, and (6) calculation of maturity levels. Access was not available for two studies. A study was included in the detailed analysis if it met more than three of the criteria. Ultimately, ten relevant studies were selected for extracting critical information to inform the development of our model (Figure 3).

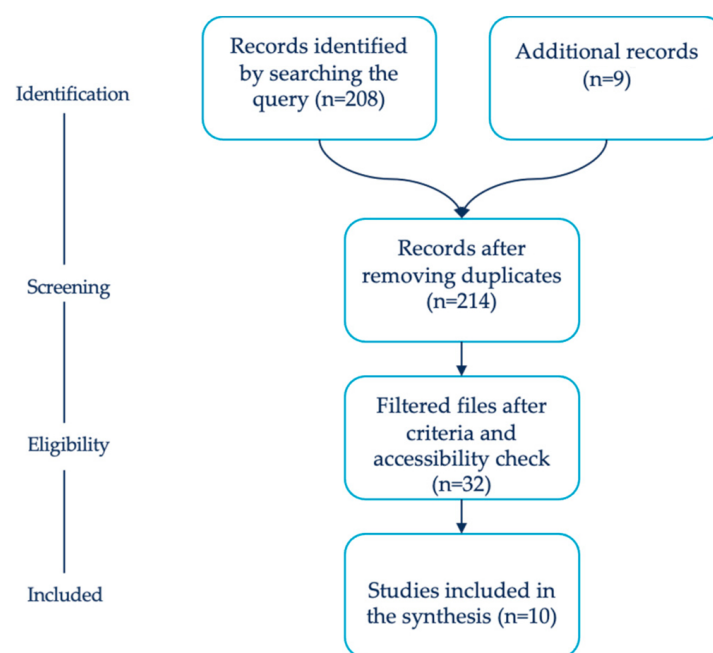


Figure 3. Systematic review procedure.

Table 2 presents the ten relevant studies and their classification according to the analyzed criteria. A study is listed in the corresponding column if it addresses a particular criterion. As the literature review revealed, most existing studies measure maturity using capabilities. Some studies use questionnaires as a basis for maturity assessments. However, other aspects, such as the calculation of maturity levels (i.e., the connection between capabilities and the final maturity level), are often neglected and not described transparently in detail. Moreover, criteria such as the analysis of as-is and to-be statuses, the development of a roadmap with action recommendations, and the use of assessment tools as well as the calculation of maturity levels are seldom addressed in research.

Table 2. Literature review.

Maturity Measurement	Use of Capabilities	The Concept for the Analysis of As-Is and To-Be Status	Development of a Roadmap for Recommendations for Action	Use of an Assessment Tool	Maturity Assessment Based on a Questionnaire	Calculation of the Maturity Level
[16,18–24]	[16,18–24]	[18,23]	[8,19,22]	[18,20]	[18–23]	[16,18,23]

These findings highlight the need for a standardized and holistic maturity model that includes a questionnaire and method of calculation to obtain the maturity level of each process dimension. Such a model should be domain-specific to data management and addressing the challenges of the industry. Therefore, the model of this work includes components such as maturity measurement, use of capabilities, maturity assessment based on a questionnaire, and calculation of maturity levels while also analyzing potential correlations between capabilities. The description of the concept of the analysis of the as-is and to-be status as well as the development of a roadmap for recommendation for action and the explicit use of an assessment will be part of the maintain phase in future work.

To focus first on the maturity map with its capabilities and the questionnaire to determine the maturity level of each capability, it is necessary to understand which capabilities are important regarding data management. Table A1 in the Appendix A presents the maturity models from the analyzed studies that incorporate process dimensions and capabilities, like those used in our reference model. Some models distinguish between the terms “process dimensions” and “capabilities”. For others, the allocation of these terms was made according to the following principle: more general terms were categorized as “process dimensions”, while more specific terms were classified as “capabilities”.

The results offer an initial foundation for comparing the refined maturity model, which was revised based on expert opinions and practical examples. This comparison helps determine whether the relevant process dimensions and capabilities identified in research are also covered by the practical approach or if further adaptations are needed. The reference model and the adaptations based on the literature review are described in the following Sections 3.2.2 and 3.2.3.

### 3.2.2. Maturity Model Structure and Content

To determine the model structure of the refined model, it is necessary to understand the structure of the reference model. The initial maturity model is based on six pillars, with twenty-six capabilities in total that more precisely elaborate the process dimensions (Figure 2).

The strategy dimension covers topics such as the data management vision, roadmap, communication, and change management. The processes dimension includes an assessment of the current state of data management processes, their architecture, technical specifications, and performance measurement.

The organization dimension addresses roles and responsibilities, governance structures, and the connection between local and global organizations. In the context of master data, it evaluates data and information models, nomenclature, semantics, and the integration of other systems. The architecture dimension analyzes data applications, the organization’s data architecture design, and standardized procedures.

Operations and support completes the data management framework by covering topics such as professional and technical support, data quality monitoring, and training and knowledge management. Each process dimension contains specific capabilities, which are measured to determine the maturity level of that dimension.

### 3.2.3. Definition of Dimensions and Capabilities

Following the initial literature review and the comparison of the data management maturity model (as reference model) and the digital transformation capability maturity model discussed in Pörtner et al. 2022 [13], it became clear that the scope of the maturity model needed to be expanded from master data to encompass overall data management. This expansion affects the process dimensions, as one of the dimensions was specifically focused on master data. Six consultant experts from the area of data management re-designed an initial draft of the refined maturity map based on project experiences in that domain.

As a result, the master data dimension was renamed as the data dimension.

Additionally, the operation and support dimension was expanded to incorporate the topic of innovation due to the increasing presence of new technologies in the digital age. Figure 4 illustrates the refined maturity model, highlighting these identified areas of improvement.

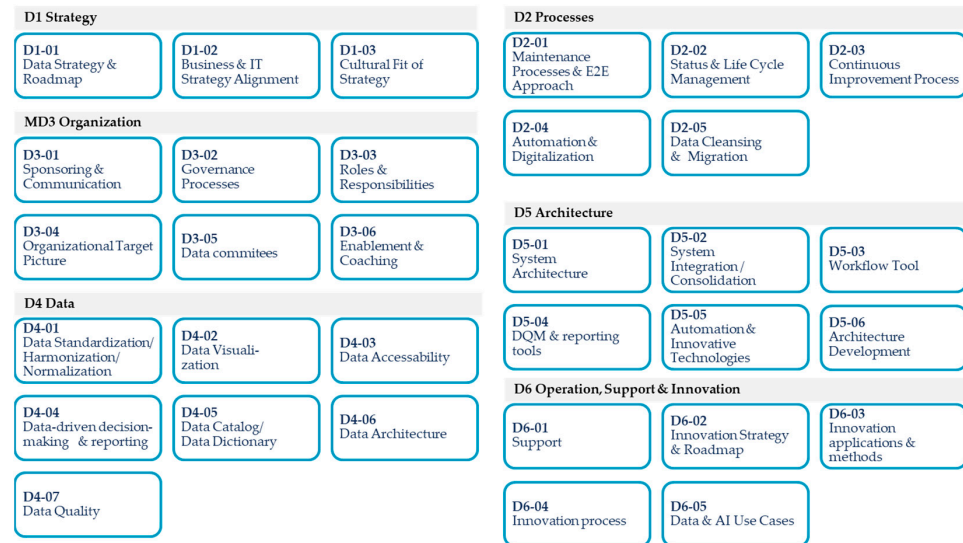


Figure 4. Design of the maturity map (based on [9]).

The following section briefly describes the content of each process dimension.

The strategy dimension encompasses all strategic instruments required for the successful implementation of data management. It also covers data strategy and its influence on both the organization’s strategic decisions and day-to-day operations.

The processes dimension includes the operational maintenance of processes, lifecycle management, and the degree of process automation. The organization dimension addresses the structure of the data organization, including roles, responsibilities, governance, standards, and guidelines for effective data management. The data dimension deals with the definition of data models, objects, their semantics, and nomenclature. This is not limited to master data but applies to various data types. The IT architecture dimension covers the technological implementation and structure of the organization’s solution architecture. The operations, support, and innovation dimension covers not only support topics but also innovation, particularly focusing on the degree of innovation and the use of emerging technologies. This is increasingly relevant in the digital age with the rapid development of new technologies.

The maturity model aims to address the key elements of a data-driven organization, including digital transformation, data science, data-driven business models, data-driven innovation, and data analytics. The strategy process dimension addresses digital transformation by focusing on corporate strategy and culture. Data science, data-driven innovation, and data-driven business models are covered by content such as data strategy and roadmap, enablement and coaching, innovation strategy and roadmap, and the organizational target picture. Data analytics is included under capability enablement and coaching.

The literature review (see Table A1) shows that the strategy dimension is explicitly mentioned in maturity models such as EIM, CMMI, DCAM, Oracle, and digital maturity (based on SCP and DCV). It is also indirectly referenced through categories such as enterprise intent (DMMM), goals and principles (DAMA-DMBOK), and strategic alignment (DDPCDM). In the refined model, we also use the term “cultural fit of strategy” to encompass both strategic content and its link to organizational and people culture.

The processes dimension is explicitly mentioned in EIM, while other models such as DMMM (data operations), Oracle (data maintenance), MD3M (maintenance), and DREAMY

(Workflow) also cover process-related content. Existing models often describe process activities in general terms, while our model includes not only process maintenance but also lifecycle management, improvement processes, automation, migration, and cleansing activities.

The organization dimension is frequently represented as data governance in the literature (e.g., DMMM, DAMA-DMBOK, CMMI, DCAM, and Dataflux). Other relevant capabilities found in the literature include sponsorship and portfolio management, organizational learning (DDPCDM), roles and responsibilities (DAMA-DMBOK), and culture (digital maturity model based on SCP and DCV). In this dimension, our model aligns well with the most cited capabilities.

The data dimension is addressed in models such as DMMM and DREAMY (data management). Data quality is explicitly mentioned in DAMA-DMBOK, CMMI, DCAM, IBM DGMM, Stanford DGMM, and MD3M. Oracle also addresses topics such as data utilization. Our model focuses on more specific capabilities such as data accessibility, data harmonization, data visualization, and data-driven decision making.

The IT architecture dimension is explicitly mentioned in DAMA-DMBOK, CMMI, DCAM, and IBM DGMM. Other models include systems (DMMM), infrastructure (EIM), integration (Dataflux), and technology (DREAMY, digital maturity model based on SCP and DCV, and data-driven organization maturity model). Our model also adds workflow tools to measure technological implementation.

The operations, support, and innovation dimension is covered by data operations in DMMM, CMMI, DCAM, and the digital maturity model (based on SCP and DCV). CMMI also includes supporting processes, but innovation and AI use cases are not extensively covered in these models. Our model explicitly addresses data and AI use cases and innovation processes.

The results of the literature review confirm that the six renamed pillars are central areas for determining the maturity level of data projects. However, many existing models lack a detailed representation of the various capabilities and the methods for calculating the maturity of each process dimension. For this reason, this paper attempts to identify possible correlations between the process dimensions and capabilities (Sections 3.3.1 and 3.3.2).

Overall, it can be stated that the refined model is close to the core principles of the CMMI data management maturity model and supplements these with relevant capabilities (Table 3). While the data management maturity model (CMMI) addresses the capabilities of data management strategy, communication, data management function, business case, and funding in the data management strategy process dimension, the model presented focuses on the strategy and roadmap; the alignment between business and IT strategy, which includes topics such as business case and funding but does not explicitly list them; and the cultural fit when it comes to the topic of a data culture. Communication and funding are assigned to the organization dimension.

The organization dimension deals more specifically with the topic of governance and touches not only on processes but also explicitly on roles and responsibilities, data committees, and the general organizational vision. Employees and the topic of enablement and coaching are also addressed. While the CMMI model in the data dimension focuses very strongly on data quality, including the strategy, profiling, and cleansing of data, the proposed model addresses data visualization, data coherency, data accessibility, data storage in terms of data cataloging, and the data architecture picture. In addition, the handling of data is evaluated regarding data-driven decision making. The topic of data cleansing with associated migration is categorized in the processes dimension of the proposed model. The processes dimension covers data operations topics from the CMMI model when it comes to life cycle management and supplements this with process maintenance topics, continuous improvement process, and innovative topics such as automation and digitalization of

processes. Both models are similar in the architecture dimension, with the difference that the proposed model also looks at automation and innovative technologies. In the support dimension, the CMMI model focuses on risk and configuration management and deals with topics such as measurement and analyses, process management, and process quality assurance. The proposed model focuses more on innovative solutions and includes capabilities such as innovation strategy and roadmap, innovation applications and methods, innovation process, and data and AI use cases.

**Table 3.** Differences and additional components of proposed model.

<b>Data Management Maturity Model (CMMI)</b>	<b>Proposed Model</b>	<b>Differences/Additional Complements of Proposed Model</b>
<b>Data Management Strategy</b> Data Management Strategy Communications Data Management Function Business Case Funding	<b>Strategy</b> Data strategy and roadmap Business and IT strategy alignment Cultural fit of strategy	Communications and Funding are included in the organization dimension.
<b>Data Governance</b> Governance Management Business Glossary Metadata Management	<b>Organization</b> Sponsoring and communication Governance processes Roles and responsibilities Organizational target picture Data committees Enablement and coaching	The organization dimension not only tackles the topic of data governance with its processes but also the organizational target picture, structures, and skillset of people (enablement and coaching).
<b>Data Quality</b> Data Quality Strategy Data Profiling Data Quality Assessment Data Cleansing	<b>Data</b> Data standardization/ harmonization/normalization Data visualization Data accessibility Data-driven decision-making Data catalog/ data dictionary Data architecture Data quality	The data dimension not only deals with data quality per se but also covers topics ranging from data access, modelling, visualization, and storage to data-driven decision-making skills.
<b>Data Operations</b> Data Requirements Definition Data Lifecycle Management Provider Management	<b>Processes</b> Maintenance processes and E2E approach Status and lifecycle management Continuous improvement process Automatization and digitalization Data cleansing and migration	The processes dimension deals precisely with maintenance processes and the life cycle management with its improvement process. In addition, data cleansing and migration are prescribed in this dimension. Automation and digitalization topics are also covered here.
<b>Platform and Architecture</b> Architectural Approach Architectural Standards Data Management Platform Data Integration Historical Data, Archiving, and Retention	<b>Architecture</b> System architecture System integration/consolidation Workflow tool DQM and reporting tools Automation and innovative technologies Architecture development	The architecture dimension covers additionally innovative topics such as automation and innovative technologies.
<b>Supporting Processes</b> Measurement and Analyses Process Management Process Quality Assurance Risk Management Configuration Management	<b>Operations, Support, and Innovation</b> Support Innovation strategy and roadmap Innovation applications and methods Innovation process Data and AI use cases	The operations, support, and innovation dimension has a strong focus on innovative topics and data and AI use cases. It does not specifically cover risk, process, or configuration management.

Terms in bold represent the different process dimensions.

In general, it can be said that innovative capabilities have been added to the CMMI model and that the categories attempt to map the status quo of the dimension in more detail with more precise capabilities.

In the model, each dimension is evaluated, and its maturity is measured based on responses to a questionnaire linked to answers which relate to a specific maturity level. The defined maturity levels of the refined model are based on the ISO/IEC 33002 standard [25].

### 3.2.4. Definition of Maturity Levels

The proposed maturity model comprises six process dimensions, with thirty-two defined capabilities used to assess overall maturity in the transition to a data-driven organization. Capability levels were based on the ISO/IEC 33002 standard [25], which provides a

framework for measuring maturity across each process dimension. The following six levels were defined: incomplete, performed, managed, established, predictable, and innovating.

Additionally, a questionnaire was developed for each capability.

The questions and answers were derived by the six consultant experts after setting up and comparing the initial maturity map draft with the literature results. To validate the feasibility and comprehensibility of this questionnaire, a short assessment was carried out in the deploy phase, in which questions were selected as examples from each process dimension, and initial feedback was obtained so that the overall questionnaire can be validated in future research in the form of an assessment in various industry sectors. This questionnaire outlines the organization’s status based on interviews with selected stakeholders, ranging from operational staff to management. To ensure objectivity and transparency, the questionnaire consists of fixed-answer options linked to each maturity level. While answering questions or providing specific examples requires some knowledge from the assessors, the assessment can be managed without extensive prior experience.

As an example, the data strategy and roadmap capability is illustrated in Table 4, which presents questions and possible answers. Six possible answers are provided according to the defined maturity levels.

**Table 4.** Questionnaire example of the strategy dimension.

#	Question	Level	Answer Option Depending on the Level
1	How would you describe your data strategy?	1	Incomplete: The organization lacks a defined or planned data strategy for the foreseeable future.
		2	Performed: An initiative is underway to establish a data strategy, albeit without a clearly defined roadmap.
		3	Managed: Efforts are underway to develop a comprehensive data strategy roadmap by conducting an assessment of the current data landscape.
		4	Established: A robust data strategy has been formulated, encompassing a detailed roadmap and specific action items to lay the groundwork for future initiatives.
		5	Predictable: The data strategy undergoes yearly review and refinement as part of an ongoing process of continuous improvement.
		6	Innovating: The data strategy is implemented continuously and has a profound, long-lasting impact on the entire organization’s value-creation efforts.
2	To what extent are the business, data, and strategy aligned?	1	Incomplete: The absence of a communicated data strategy leaves stakeholders uninformed.
		2	Performed: Plans are in place to align data, business, and IT strategies, though communication remains pending.
		3	Managed: Alignment between data and business strategies has commenced, accompanied by targeted communication efforts.
		4	Established: The current data strategy aligns with the overarching business strategy, with effective communication channels established.
		5	Predictable: Data considerations are seamlessly integrated into the annual company-wide strategy review and adjustment processes.
		6	Innovating: The overarching data strategy actively facilitates and enhances alignment with the company’s business strategy through continuous communication and management.
3	To what degree do the data (and business) strategy fit the company culture?	1	Incomplete: The prevailing company culture does not accommodate a data strategy.
		2	Performed: Plans are underway to align the future data strategy with the organization’s cultural ethos.
		3	Managed: Initial efforts to integrate company culture within the data strategy framework are in progress.
		4	Established: The data strategy is harmonized with the existing cultural fabric of the organization.
		5	Predictable: Cultural alignment forms an integral part of the annual strategy review, ensuring cohesion between data strategy and organizational culture.
		6	Innovating: Ongoing strategy reviews actively foster a supportive data culture within the company, driving innovation and alignment.

### 3.3. Phase 4—Test

The refined maturity model was tested for validity and relevance by performing an industry survey to collect challenges from the industry and to calculate potential correlations.

#### 3.3.1. Industry Survey

To identify the most challenging process dimensions and capabilities that need to be addressed, an industry survey on the topic of “The Value of Data-Driven Organizations” was conducted. Additionally, to explore potential correlations between capabilities, the analytic hierarchy process (AHP) method was employed with input from six consultant experts in data management. To calculate the correlations within a process dimension,

we utilized the decision support method of multi-criteria (AHP), which is a decision support method to support decisions in teams and make the decision making and the result comprehensible. It also tries to uncover possible inconsistencies in decision making [26,27].

The survey included a questionnaire that addresses topics such as strategy, governance, processes, architecture, data literacy, and data quality. Furthermore, questions about digitalization and economic success were included to provide an initial overview of the value associated with data-driven organizations. In total, 34 questions were designed to cover these topics.

For the 34 questions, a distinction was made between variable types: moderator and variable. Ordinal scales were utilized with three scale characteristics (yes/no, a one-to-seven rating scale, and open-ended questions) to enhance comparability. The one-to-seven rating scale was specifically chosen to discourage respondents from consistently selecting the midpoint, thereby ensuring more precise results.

The developed questions were initially tested by eleven experienced data and analytics consultants from Camelot, focusing on criteria such as comprehensibility, completeness, and usefulness. Feedback from this review was incorporated to improve the feasibility of the study. The online survey was distributed to industrial customers from May to September 2023. The complete list of questions and answer options can be found in the Appendix A (see Table A2).

Responses from 20 participants formed the population N for the data evaluation. The survey was qualitative and targeted, as it was aimed at proven experts from various industry sectors. The participants represented specific, practical knowledge that could not be equally covered by a larger, less specialized sample. This broad distribution created a representative basis for cross-industry statements. The number of participants was chosen so that a saturation point could be reached at which opinions and findings would begin to repeat themselves, which would indicate that the relevant perspectives were sufficiently covered.

Initially, the general questions (questions 1–5) were analyzed to obtain an overview of the key characteristics of the organizations involved.

Most respondents (40 percent) represented organizations from the automotive sector, while other sectors included travel (10 percent), pharma/life sciences (5 percent), consumer goods/manufacturers (10 percent), and others (35 percent) (see Figure 5).

Which industry does your company belong to?

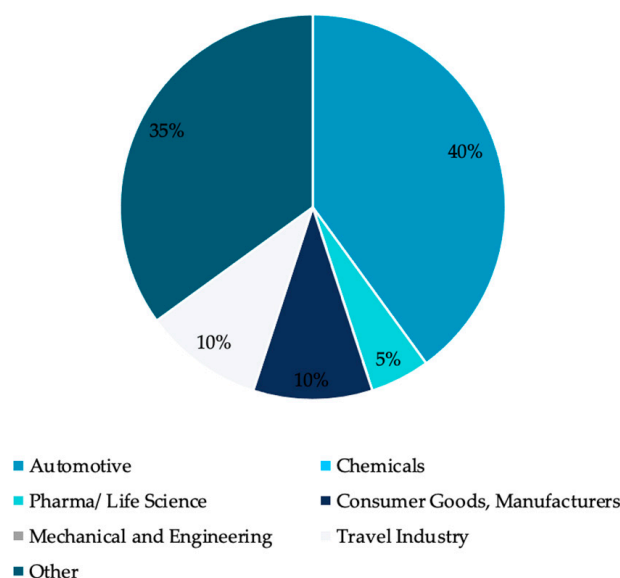


Figure 5. Industry sectors of participants.

Overall, 55 percent of the organizations had a revenue of more than 5001 million euros. As far as the number of employees of the organizations is concerned, both small- to medium-sized enterprises (up to 5000 employees) were represented with 35 percent as well as large companies (>100,000 employees) with 25 percent.

When asked about the existence of a data strategy in the organization, 75 percent mentioned having one.

Furthermore, 40 percent saw the added value of the organization’s data strategy about business goals to be at least above average or higher. This shows how important it is to have a well-defined data strategy.

Only 25 percent stated that roles and responsibilities are strongly to very strongly implemented in the organization. This points to a challenge in governance.

While 11 out of 20 indicated that data processes are relatively or better implemented in the organization, there is a need for automation in processes, as only 7 out of 20 answers indicated its relative or better implementation.

In general, more than half of all respondents rated their company business performance as relatively strong in the last year.

In the next step, correlations between answers to the topic-specific questions 7 to 33 were investigated for making better roadmap improvement decisions. To do so, the Spearman’s correlation coefficient was used. Spearman’s correlation coefficient calculates the strength of the relationship between two variables ( $\rho$ ) on ordinal scales. The correlation coefficient is defined between  $-1$  and  $+1$ . A value of  $+1$  corresponds to a perfect positive relationship between both variables, while a correlation of  $-1$  describes a perfect negative relationship. A value of zero indicates that there is no relationship between the two variables. According to Cohen’s guidelines, a distinction can be made between very low ( $|\rho| < 0.10$ ), low ( $|\rho| = 0.10-0.30$ ), medium ( $|\rho| = 0.30-0.50$ ), and high ( $|\rho| > 0.50$ ) correlation. Statistical significance ( $p$ -value) can then be used to check how likely it is that a result is based on chance. The significance level was set at  $\alpha = 0.05$ . For this purpose, the  $p$ -value was calculated accordingly to determine the credibility of the statements. If the  $p$ -value was less than 0.05, significance was given.

Selected correlations across the dimensions about investment in digitalization were determined. The correlations are shown in the Table 5 below. Correlations that are underlined represent a weak positive/negative linear relationship, and correlations marked in bold represent a strong one. All other correlations are medium correlations.

**Table 5.** Calculated Spearman-Rho correlations.

	Innovative Power (Q29)	Data Processes (Q12)	Future Competitiveness (Q28)	Data-Literate Organization (Q23)	New Business Models (Q24)	Automated Processes (Q14)	Invest Data Architecture	Employee Trained (Q20)	Gut Feelings (Q19)	
N = 20 Spearman-Rho *	Invest in Digitalization (Q32)	0.316	0.431	0.432	<u>0.029</u>	<b>0.523</b>	0.279	0.423	0.326	<b><u>-0.595</u></b>

\* Significant at the 0.05 level.

Organizations that are increasingly investing in digitalization see a higher innovative power in their organization ( $\rho = 0.316$ ,  $p$ -value = 0.174). However, with a  $p$ -value of 0.174, this correlation is not statistically significant, implying that although there may be a relationship, we cannot be confident in its existence without a larger sample or more rigorous analysis.

In addition, investments are made in areas such as the implementation of data processes ( $\rho = 0.431$ ,  $p$ -value = 0.058), data architecture ( $\rho = 0.423$ ,  $p$ -value = 0.063), and automation of processes ( $\rho = 0.279$ ,  $p$ -value = 0.234). The positive correlations between

investment in digitalization and both the automation of processes ( $\rho = 0.279$ ) and data architecture ( $\rho = 0.423$ ) show some potential relationships, though again, the  $p$ -values suggest these correlations are not strong enough to draw definitive conclusions.

Furthermore, a weakly positive, linear correlation ( $\rho = 0.326$ ,  $p$ -value = 0.161) between investment in digitalization and trained employees in the field of data and analytics skills was determined. Although the results calculated a positive correlation between the calculated variables  $x$  and  $y$ , some correlations are not statistically significant at the usual significance level of 0.05 due to the higher  $p$ -value.

The most significant positive correlation ( $\rho = 0.523$ ,  $p$ -value = 0.018) is between investment in digitalization and the development of new business models, suggesting a notable connection. This result indicates that organizations investing in digital technologies are more likely to explore and develop innovative business approaches.

Moreover, the investment in digitalization is positively correlated with future competitiveness ( $\rho = 0.432$ ,  $p$ -value = 0.057). The need for a data-literate organizations regarding digitalization is not yet recognizable due to a very weak correlation of 0.029 ( $p$ -value = 0.903). There is little evidence that changes in one variable are related to changes in the other. The high  $p$ -value of 0.903 indicates that the data do not provide sufficient evidence that true correlation exists.

A negative correlation ( $\rho = -0.595$ ,  $p$ -value = 0.006) exists between digital investments and decisions based on gut feelings. This significant finding suggests that as companies invest in digital capabilities, reliance on instinct or informal decision-making decreases, which supports the view that digitalization fosters more data-driven approaches.

Surprisingly, the correlation between digitalization and data-literate organizations is almost non-existent ( $\rho = 0.029$ ,  $p$ -value = 0.903). This suggests that current investments in digitalization are not yet translating into increased data literacy within organizations. Given the high  $p$ -value, this result is likely due to randomness, pointing to a disconnect between digital investment and the development of a data-savvy workforce.

Overall, it can be said that the determined correlations do not prove causality in some cases. Nonetheless, a certain trend can be discerned from the results, which is consistent with the overall impression on the topic of data management and is also reflected in the challenges mentioned: digitalization is recognized as a relevant topic in companies, and investments related to managing and dealing with data are made to stay competitive in the future.

In the next step, one representative question from the survey was chosen for each process dimension to calculate the overall correlation between the process dimensions (Table 6). These correlations can be seen as a first indicator that can be used later for implementation purposes while building up a roadmap for maturity improvement.

**Table 6.** Correlation of process dimensions.

	Strategy (Q9)	Processes (Q12)	Organization (Q11)	Data (Q26)	Architecture (Q15)	Operations, Support, and Innovation (Q29)
<b>Strategy (Q9)</b>		<b>0.61</b> <i>p</i> -value 0.004	<b>0.648</b> <i>p</i> -value 0.001	0.299 <i>p</i> -value 0.201	<b>0.739</b> <i>p</i> -value 0.0002	0.274 <i>p</i> -value 0.243
<b>Processes (Q12)</b>	0.61		<b>0.719</b> <i>p</i> -value 0.0004	0.265 <i>p</i> -value 0.259	<b>0.746</b> <i>p</i> -value 0.0002	0.016 <i>p</i> -value 0.946
<b>Organization (Q11)</b>	0.648	0.719		0.44 <i>p</i> -value 0.052	<b>0.759</b> <i>p</i> -value 0.0001	0.235 <i>p</i> -value 0.318
<b>Data (Q26)</b>	0.299	0.265	0.44		0.450 <i>p</i> -value 0.046	<b>0.651</b> <i>p</i> -value 0.002
<b>Architecture (Q15)</b>	0.739	0.746	0.759	0.450		0.335 <i>p</i> -value 0.149
<b>Operations, Support, and Innovation (Q29)</b>	0.274	0.016	0.235	0.651	0.335	

Numbers showing a strong positive linear correlation are highlighted in bold. The dimension of strategy has strong connections to the process dimensions of processes, organization, and architecture. The dimension of processes has strong connections to the dimensions of strategy, organization, and architecture. The dimension of data is only strongly correlated with the dimensions of operations, support, and innovation. In general, rather weak/medium correlations (<0.5) of the dimensions of data and of operations, support, and innovation with the other process dimensions are striking.

The four process dimensions with the most significant correlations—strategy, processes, organization, and architecture—suggest that improvement in one dimension (e.g., strategy) could lead to positive outcomes in the others, indicating a multiplier effect. Focusing on these dimensions could have the most substantial impact on an organization’s overall maturity.

The positive correlation between digitalization and the development of new business models indicates that investments in digital technologies can unlock new revenue streams and business innovations, reinforcing the strategic importance of digitalization.

The weak correlation between digital investments and data literacy raises concerns that while organizations are investing in technologies, they are not sufficiently developing the human capital required to maximize the potential of these technologies. Future investments should prioritize building data literacy alongside technological infrastructure.

In conclusion, while the survey provides valuable insights into the process dimensions that influence organizational maturity, its methodological limitations warrant cautious interpretation of the findings. The Spearman correlation analysis highlights relationships between variables (e.g., investment in digitalization and innovative power), but it does not establish causality. While some significant correlations were found, these do not imply that one variable directly causes changes in another. The absence of longitudinal data prevents causal claims about how investments influence organizational performance over time. Further studies in the future with larger samples and more rigorous statistical methods are needed to confirm these relationships and inform effective strategies for data-driven transformation.

### 3.3.2. Evaluation of Process Dimensions and Maturity Levels

Since the different process dimensions have different influences on each other, and not all dimensions can be tackled at the same time after determining an overall maturity level, it is advantageous to understand if correlations exist and to capture these relationships when weighting the maturity level calculation. For this purpose, the expert knowledge from six data management consultants was used. The following weightings and their feasibility are illustrated using two process dimensions as examples. An initial validation of the correlations and the feasibility of the questionnaire structure was performed in the form of a short assessment in the deploy phase.

The six process dimensions are included in the overall maturity level with equal proportions. To calculate the correlations within a process dimension, the AHP method was used.

First, an evaluation matrix was set up for each process dimension. The calculation of the weighting of capabilities within the process dimension strategy is used as an example. The capabilities are compared in rows and columns, which allows for a pairwise evaluation. This evaluates whether the row is more important than the column.

The rating scale (1 to 9) used was Saaty’s scale, which measures the importance or difference of one element from another. The diagonal contains a “1” in each case, as these are only the criteria against themselves. This applies to the elements of the matrix:

$$a_{j,i} = \frac{1}{a_{i,j}} \quad (1)$$

After a complete, independent evaluation by six experts in the field of data management, the weighting of each criterion was calculated. First, the average of all opinions was

considered to continue the further calculation with this result. An example of a pairwise comparison matrix is shown in Table A3.

Eigenvalues (EVs) were required for the evaluation. This means that the criteria could be evaluated as independently of each other as possible. Because six experts filled in the pairwise comparison matrix, the geometric mean of all ratings was used for further analysis. Table A4 in the annex contains the matrix with geometric mean values as an example of the process dimension of strategy.

The equation of weighting is given by the weight of each criterion and is the sum of the normalized criteria. For this purpose, the matrix is successively squared, and the eigenvector is determined. The calculation stops when the difference between two consecutive computational steps is negligible. For the calculation of the eigenvector EV, the series sum is formed and normalized. Table A5 in the annex shows the normalized matrix. From this, the EV was derived.

If we look at the eigenvector as a percentage, the criteria weights are represented in a Pareto representation (Figure 6).

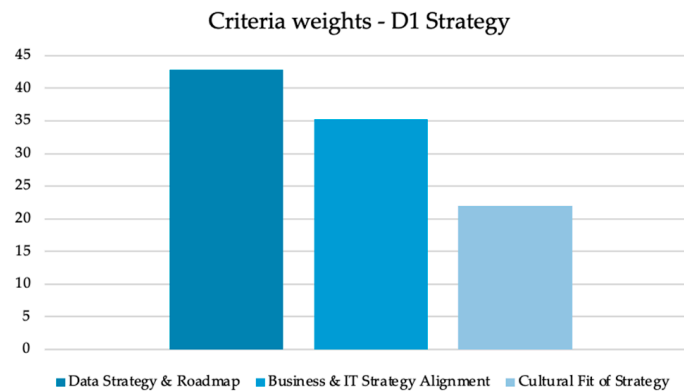


Figure 6. Eigenvector per capability of process dimension strategy.

For the assessment of inconsistencies, a consistency ratio CR is introduced, whereby  $CR < 0.1$  is considered harmless. For this purpose, the  $\lambda_{max}$  and the consistency ratio CR must first be determined.

The consistency ratio (CR) is calculated using the formula  $CR = CI/RI$ . RI is related to the dimension of the matrix and has a random consistency index from Golden and Wang [28]. The formula for calculating the consistency index CI is as follows:

$$CI = \frac{\lambda_{max} - n}{n - 1} \tag{2}$$

This calculation was performed for all process dimensions with their capabilities. The results are shown in the following Table 7.

Table 7. Calculated consistency index CI.

#	Process Dimension	CI
D1	Strategy	1.07
D2	Processes	0.75
D3	Organization	0.48
D4	Data	0.62
D5	IT Architecture	0.27
D6	Operations, Support, and Innovation	0.23

The CI values are all above and, in some cases, significantly above the permissible limit of 0.1. This indicates an inconsistency in the pair comparison judgments in the initial evaluation

matrix, which can be attributed to the subjectivity of the individuals. A group discussion and, if necessary, adjustment of individual evaluations is required to achieve the consistency criterion.

The criteria weights were then included in the fixed calculation of a process dimension maturity. If the respondent chose the following answers based on his or her possible answers, the maturity level 2.428 was obtained for the process dimension of strategy. It is calculated by multiplying the level with its criteria weight factor.

$$(3 \times 0.428) + (2 \times 0.352) + (2 \times 0.219). \quad (3)$$

A maturity level calculated without any criteria weights would be 2.333.

$$((3 + 2 + 2)/3). \quad (4)$$

Moreover, the process dimensions all have the same influence on determining the overall maturity level. No further weightings are given here. The extent to which the defined weightings are valid has not yet been checked in the form of validation; this will be checked in the future.

However, the calculated and analyzed correlations between all process dimensions have an influence on the implementation roadmap of improving the maturity of process dimensions. As analyzed in the challenges and by the evaluation of the correlation between different process dimensions, it can be stated that the most influencing process dimensions are the strategy, processes, organizations, and architecture dimensions, while the biggest challenges occur in the strategy and organization dimensions.

### 3.4. Phase 5—Deploy

The model was further validated through a short assessment to test for generalizability and collect feedback for improvements.

#### 3.4.1. Short Assessment

The deployment of the refined model, along with its questionnaire, was conducted as a short assessment since the full questionnaire was not tested but rather only representative questions involving 31 experts from various industry sectors with expertise in data and analytics. The participants were a group of experts from various industry sectors. This allowed for more detailed, well-founded insights that carried more weight than the results of less specialized participants. Representativeness was achieved through the quality and relevance of the opinions collected. During this short assessment, participants were asked to explicitly identify the most significant current challenges. These challenges were then compared to the initial challenges identified in the earlier survey results to confirm the selection of dimensions and capabilities.

A representative number of questions were chosen for each process dimension to provide an initial overview of the organization's maturity and to validate the refined maturity levels, dimensions, and the questionnaire in this first iteration.

Figure 7 illustrates the distribution of participants by industrial sector. The majority of participants were from the automotive sector (31 percent), while the chemicals and pharma sectors collectively accounted for the second-largest group, with 5 out of 31 participants.

The results of the short assessment showed that the top-performing organization had a maturity level of 4.83 out of 6, while the average had a maturity level of 3.61, which is between a managed and established maturity level. In the following Figure 8, the average maturity levels of the six process dimensions are presented (regarding all industries and best in class). Best in class is always the value of the highest maturity in each process dimension. Each best-in-class value can be from a different organization. Additionally, the average of the chemicals industry is mentioned, as this industry showed the best results compared to all industry sectors. The result here reflects a trend and is intended to serve as an indication, as the number of participants in the various industries varied.

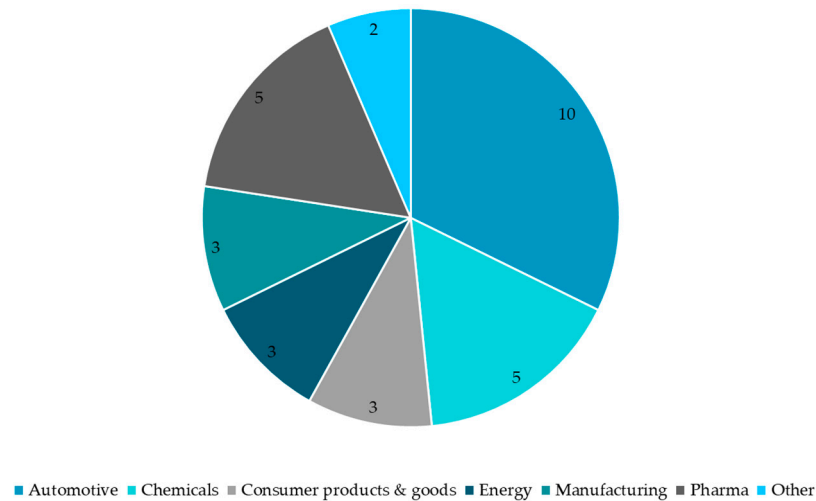


Figure 7. Participants' industries.

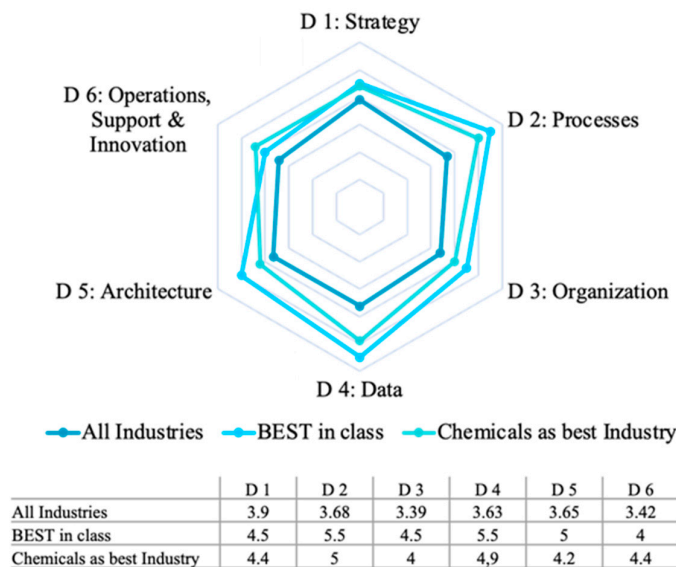


Figure 8. Results of short assessment per process dimension.

The dimensions of organization (average score of 3.39) and operations, support, and innovation (average score of 3.42) exhibited lower degrees of maturity when comparing the average values across all industries. Overall, the maturity levels identified through the short assessment primarily fell between levels 3 and 4. This indicates that, moving forward, organizations must continue to recognize the value of data and utilize them not only to advance technology in line with Industry 5.0 but also to address existing organizational shortcomings. This reinforces findings from the correlation analysis, which suggest that while investments are being made in digitalization, most companies are not yet fully capitalizing on the added value of data.

Table 8 presents the challenges identified in the short assessment. Moreover, all identified challenges are mapped to the corresponding capabilities that address them within our model. Notably, no challenges were assigned to the operations, support, and innovation areas. This is likely because the foundational elements are not yet fully developed, making innovative topics a lower priority or not yet relevant. However, the model is flexible and can be improved or extended as new challenges arise. When this occurs, a recalculation of weightings should be performed to reflect the current situation. The most significant challenges are associated with the dimensions of strategy and organization. While the

maturity of the strategy process dimension appears to have the highest overall average maturity across all industries, the organization dimension still has considerable room for improvement. The same challenges arise regardless of the industry sector. It is therefore crucial to specifically address these challenges in the maturity model to be able to tackle them in the next step when increasing maturity.

**Table 8.** Challenges in process dimensions.

Dimensions	Challenges	Capabilities
D1 Strategy	Aligned strategy becoming a data-driven organization Missing end-to-end strategy The value of data is not recognized widely Overcome thinking in silos	Data strategy and roadmap
	Aligned strategy becoming a data-driven organization	Business and IT strategy alignment
	Data culture not established Aligned strategy becoming a data-driven organization	Cultural fit of strategy
D2 Processes	End-to-end cross-platform business processes	Maintenance processes and E2E approach
		Status and life cycle management
		Continuous improvement process
	No automation Data cleansing	Automation and digitalization Data cleansing and migration
D3 Organization	Data culture not established The value of data is not recognized widely Business empowerment	Sponsoring and communication
	Governance structures Overcome thinking in silos	Governance processes
	Data ownership Governance structures Absence of data literacy on an organizational level	Roles and responsibilities
	Governance structures Overcome thinking in silos	Organizational target picture
	Governance structures	Data committees
	Data culture not established Absence of data literacy on organizational level Business empowerment Quality Overcome thinking in silos	Enablement and coaching
D4 Data	Data harmonization	Data harmonization
		Data visualization
	Data distribution across units Data access	Data accessibility
		Data-driven decision making and reporting
		Data dictionary
	Data distribution across units Overcome thinking in silos Quality	Data architecture Data quality
D5 Architecture	Integration Interfaces Heterogenous application landscape Overcome thinking in silos	System architecture
	Integration	System integration/consolidation
		Workflow tool
		Data quality management and reporting tools
		Automation and innovative technologies
D6 Operations, Support, and Innovation		Architecture development
		Support
		Innovation strategy and roadmap
		Innovation applications and methods
	Innovation process	
	Data and AI use cases	

### 3.4.2. Feedback Gathering

Following discussions with participants from the short assessment, it became evident that the feasibility and comprehensibility of the questions were satisfactory, ensuring a transparent maturity assessment. Feedback indicated that it is crucial to clearly define the scope of the questions at the outset, as significant local variations often exist compared to a global organizational perspective on what should be evaluated.

Additionally, we recognized the importance of selecting the right participants for the assessment. Interviewing both operational staff and management is essential, as this approach provides a more valid and comprehensive view of the assessment's scope.

### 3.4.3. Evaluation of Results

The results of the short assessment provide an initial overview of where organizations currently stand. The average maturity level across all industries indicates that the "short assessed" organizations have maturity levels between level 3, i.e., managed, and level 4, i.e., established. To maximize the return on investments in digitalization, it is crucial to understand what needs to be achieved to reach level 4, i.e., established.

The level definitions are based on ISO 33002 and were supplemented by the project experience of data management experts. It is important to note that the ISO standard levels start at level 0, i.e., incomplete, while this approach begins at level 1, i.e., incomplete.

An overall maturity level of 1 indicates that there is no management focus on data topics and no existing vision or strategy regarding data and analytics. In contrast, level 2, i.e., performed, signifies that while there is no overarching vision or strategy, data issues are addressed as needed. At this level, there is silo optimization in data management, and no single point of maintenance is defined. Furthermore, no key performance indicators (KPIs) or reporting mechanisms are established or utilized.

At level 3, i.e., managed, where the average of the "short assessed" organizations is located, a vision and strategy are defined. Roles and responsibilities, along with guidelines and standards, are mostly established, though they may not always be consistently followed within the organization. KPIs are defined and measured, and system integration is ensured.

To advance to level 4 (established), an organization's vision and strategy must be regularly reviewed and adapted. Organizational governance should define and ensure adherence to guidelines and standards. In addition, system integration must not only be ensured but also involve leading systems throughout their lifecycle. A continuous improvement process must also be in place.

Achieving a predictive level 5 means that management decision making is data-driven, with investments in new technologies for data management informed by data analysis. Furthermore, the functionalities of the entire organization must be integrated to enable real-time product-service and process-specific data. Quantitative performance management, data analytics, and enterprise application maintenance should be well established. The transition from a predictive to an innovative organization requires a robust data culture that aligns with management strategy and is practiced in daily operations. This also entails self-optimized decision management and business process integration throughout the entire lifecycle.

For organizations at level 3 (managed), this indicates that a fundamental foundation is in place for the transition to a data-driven organization, with a clear focus on process improvement [6]. According to the definition of a data-driven organization, "to become data-driven, not only data and technology capabilities are significant indicators but also organizational capabilities, including fostering a data culture, are vital on the journey to becoming a data-driven organization" [8]. This means that value is generated from data, as processes, systems, and structures are not only defined but also established, with a continuous improvement process actively in place. The strategy and vision should be

regularly reviewed and adapted, which is essential in the digital age with its constant influences and changes, ultimately leading to the establishment of a data culture.

For the process dimensions of the model, the strategy, processes, and organization dimensions form the foundational framework that supports the data and architecture dimensions. The operations, support, and innovation dimension becomes increasingly important for achieving higher levels (5 and 6), as data must be leveraged in new and innovative technologies, which requires a certain level of data quality. This is further illustrated by the correlation analysis of the process dimensions, showing a value of 0.651 between the dimensions of data and operations, support, and innovation. Figure 9 displays the calculated correlation strength among the different dimensions. These calculations represent a trend that should also be considered when deciding which dimensions should be tackled first if there is potential for improvement.

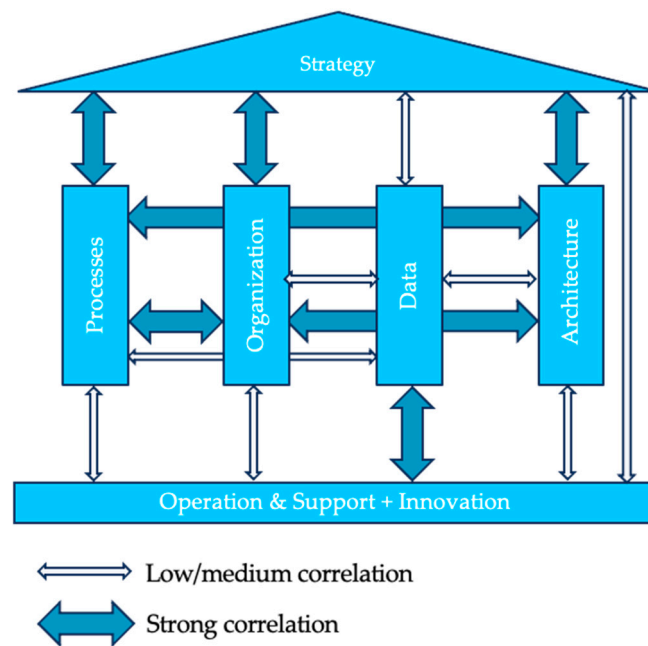


Figure 9. Correlation between process dimensions.

A trend towards strong correlations can be seen in the strategy and organization, strategy and processes, and strategy and architecture dimensions. The data strategy ultimately shapes their characteristics. The data dimension, on the other hand, predominantly has only low correlations with other dimensions. This can be explained by the fact that everything to do with the topic of data quality first requires the foundation of a strategy, processes, organization, and architecture.

### 3.5. Phase 6—Maintain

The sixth phase involves maintaining the maturity model in practice, fully validating the assessment, and analyzing deep-dive topics identified as important challenges to address in conjunction with the other phases. This phase is not yet fully completed. Preparations have included the comprehensive documentation of the model’s constructs and its online survey, which encompasses all questions designed to measure maturity. Additionally, a set-up for benchmarking has been established to track the maturity levels across different industry sectors and identify associated challenges.

This first iteration highlights the areas of focus for the maturity model as it progresses to the next iteration of development: It is evident that significant attention should be directed toward the organization dimension, particularly governance topics, in the future. Developing

a data organization with personnel who effectively manage data is crucial for advancing toward Industry 5.0. This necessity emphasizes the urgent need to address data literacy more intensively, equipping the organization and its employees to reach the next level. With this scope, the development of the maturity model will be further explored in future work.

#### 4. Discussion, Limitations, and Conclusions

Many existing maturity models lack a holistic approach applicable across industries. To ensure broad usability, capability terms must be generic, while their interpretation should be tailored to specific organizations. The model in this study was refined using a literature review, industry survey, and short assessment.

The short assessment yielded an average maturity level of 3.61, indicating a transition between “managed” and “established”. The lowest levels were recorded in the strategy and organization dimensions, highlighting challenges such as weak data culture, insufficient data literacy, and unclear roles. Despite significant digitalization investments, only 15 percent of surveyed organizations use data for decision making, indicating a gap between investment and realized benefits. An average maturity of 3.39 in the organization dimension further underscores vulnerabilities. Only 30 percent of organizations considered themselves as data-literate, and 20 percent reported having clear role definitions. These findings emphasize the need for continued digitalization and the transition to data-driven operations. The refined maturity model aims to address these gaps.

The literature review confirms the model’s relevance but highlights a lack of detail in capability levels and maturity calculations. The proposed model tries to incorporate the key aspects of data-driven organizations, including digital transformation, data science, and data analytics. However, it has some limitations.

First, the model’s scope is restricted by reliance on a reference maturity model and a focus on data management, limiting stakeholder involvement to industry and consulting experts.

Second, there are several potential confounding factors in the correlation analysis that could have influenced the results of the studies. The small number of participants reduced the statistical significance of the results. With a larger sample, more stable and reliable correlations could be found. Nevertheless, the survey ( $n = 20$ ) and short assessment ( $n = 31$ ) included only highly knowledgeable participants, ensuring opinion saturation despite a small sample size. Furthermore, 40 percent of the participants came from the automotive industry, which means that the results may not be fully transferable to other industries. Industry-specific differences could distort the correlations. In addition, a mix of small- and medium-sized companies (35 percent) and very large corporations (25 percent) could lead to a distortion, as larger companies tend to have more resources for digitalization and data management. The study only provides a first overview and cannot prove any causal relationships over time. Changes in digitalization strategies may only become noticeable after several years. External influences such as the economic situation, regulatory requirements, or technological developments were not taken into account but could have a significant influence on digitization decisions. Response bias was not analyzed in detail. Participants could have given socially desirable answers, especially regarding questions about data strategy or digitalization. The choice of scale (1–7) was intended to avoid participants often choosing the center, but they could still have tended to give moderate answers, which could influence the results.

Third, the model is at an early validation stage. The validation focuses is on capabilities rather than practical application. The short assessment provided initial feedback, but subjectivity and participant roles may have influenced results. Future validation will refine the assessment methods.

Fourth, the scalability of the short assessment remains theoretical and requires further evaluation in the maintain phase. The current weightings are based on expert input rather than on practical validation. Future iterations will involve extended industry testing in various sectors (e.g., automotive, aviation, consumer goods, pharma, and media).

Fifth, the Spearman correlation analysis identified relationships but did not establish causality. The findings serve as a basis for prioritizing improvements.

Sixth, the AHP method was used for capability weighting, but initial evaluations showed inconsistencies (CI values exceeded 0.1). Further refinements through expert discussions are necessary.

Seventh, the maintain phase is incomplete. Future efforts will refine the model through additional assessments, particularly in the organization dimension.

In conclusion, this study presents a refined maturity model for data management that is applicable across different industries. It builds on a reference model that initially focused on master data but was expanded to address general data management. The model, developed following De Bruin’s maturity assessment framework, integrates results from the literature review, an industry survey, and a short assessment.

Future work will complete the maintain phase and validate the model through extended industry testing. Maturity levels, weightings, and assessment criteria will be evaluated across multiple sectors. The focus will be on improving organizations’ ability to extract value from data by enhancing data literacy, governance, and effective data utilization.

**Author Contributions:** Conceptualization, L.P. and B.S.; methodology, L.P.; validation, L.P. and B.S.; formal analysis, A.R.; writing—original draft preparation, L.P. and B.S.; writing—review and editing, L.P., A.R. and R.M.; visualization, B.S.; supervision, R.M. and M.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Data are available on request.

**Acknowledgments:** Thanks to the co-authors Lara Pörtner, Benedikt Schmidt, Marcel Leclaire and Robert Möske, it was possible to enrich this research with practical insights and expert knowledge. Moreover, this research was partially supported by Antonia Franke, Leo Merkel, and Lukas Breuer, who provided additional focus based insight and expertise as data management consultants that greatly assisted the research.

**Conflicts of Interest:** Authors Benedikt Schmidt, Marcel Leclaire and Robert Möske were employed by the company CAMELOT Management Consultants AG. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Appendix A

**Table A1.** Maturity models from research.

Maturity Model	Process Dimension	Capabilities
Data Management Maturity Model (DMMM) [18]	Enterprise and Intent	Business Strategy, Culture, and People
	Data Management	Data Collection and Availability, Metadata Management and Data Quality, Data Storage and Preservation, Data Distribution and Consumption, Data Analytics/processing/Transformation, Data Governance, Data Monitoring and Logging
	Systems	Architecture and Infrastructure, Data Integration, Security
	Data Operations	Processes, Data Deployment, and Delivery

**Table A1.** *Cont.*

Maturity Model	Process Dimension	Capabilities
DAMA-DMBOK [29]	Data Governance	Goals and Principles Organization and Culture Activities Deliverables Roles and Responsibilities Practices and Techniques Technology
	Data Architecture Management	
	Data Development	
	Database Operations Management	
	Data Security Management	
	Reference and Master Data Management	
	Data Warehousing and Business Intelligence Management	
	Meta Data Management	
Gartner Enterprise Information Management Maturity Model (EIM) [29,30]	Vision Strategy Metrics Governance People (Org/Roles) Process (Life Cycle) Infrastructure	
Data Management Maturity Model (CMMI) [20,31]	Data Management Strategy	Data Management Strategy Communications Data Management Function Business Case Funding
	Data Governance	Governance Management Business Glossary Metadata Management
	Data Quality	Data Quality Strategy Data Profiling Data Quality Assessment Data Cleansing
	Data Operations	Data Requirements Definition Data Lifecycle Management Provider Management
	Platform and Architecture	Architectural Approach Architectural Standards Data Management Platform Data Integration Historical Data and Archiving
	Supporting Processes	Measurement and Analytics Process Management Process Quality Assurance Risk Management
	Configuration Management	
Data Management Capability Assessment Model (DCAM) [29,32]	Data Management Strategy Data Management Business Case and Funding Model Data Management Program Data Governance Data Architecture Technology Architecture Data Quality Data Operations	
IBM Data Governance Council Maturity Model (DGMM) [20,31]		Organizational Structures Stewardship Policy Value Creation Data Risk Management and Compliance Information Security and Privacy Data Architecture Data Quality Management Classification and Metadata Information Lifecycle Management Audit Information, Logging, and Reporting
Stanford Data Governance Maturity Model (Stanford DGMM) [20,31]	Awareness	People, Policies, Capabilities
	Formalization	
	Metadata	
	Stewardship	
	Data Quality	
Master Data		

**Table A1.** *Cont.*

Maturity Model	Process Dimension	Capabilities
Oracle Corporation Model [21]		Data Source Profiling Data Strategy Definition Data Consolidation Plan Definition Data Maintenance Data Utilization
Master Data Management Maturity Model (MD3M) [21,33]	Data Model	Definition of Master Data, Master Data Model, Data Landscape
	Data Quality	Assessment of Data Quality, Impact on Business, Awareness of Quality Gaps, Improvement
	Usage and Ownership	Data Usage, Data Ownership, Data Access
	Data Protection	Data Protection
Dataflux [20,31]	Maintenance	Storage, Data Lifecycle
	Architecture Governance Management Identification Integration BPM of Master Data	
Digital Readiness Assessment Maturity Model (DREAMY) [19]		Technology Internet Connectivity Leadership Equipment/Tools Cloud Storage Skill Requirement Productivity Work Environment Workflow Employee Engagement Data Management Teamwork Knowledge Sharing Decision Making Collaboration Transparency
Digital Maturity based on SCP and DCV [23]	Strategy Market Operations Culture Technology	
Data-Driven Organization Maturity Model [24]	Organization Technology Decision Process People Analytics	
Data-Drivenness Process Capability Determination Model (DDPCDM) [8]		Change Management Skill and Talent Management Strategic Alignment Sponsorship and Portfolio Management Organizational Learning

**Table A2.** Questionnaire of industry survey.

Nr.	Question	Possible Answers
1	Please enter the name of your company:	Open text answer
2	Please enter the name of your department:	Open text answer
3	If you want to receive the results of the present study, please enter your email address:	Open text answer
4	To which industry does your company belong?	Automotive
		Chemicals
		Pharma/Life Science
		Consumer Goods, Manufacturers
		Mechanical and Engineering
		Travel Industry
		Other

Table A2. Cont.

Nr.	Question	Possible Answers
5	What is the revenue of your company? (in million EUR)	<100
		100–250
		251–500
		501–1000
		1001–5000
		>5001
6	What is the size of your company? (Employees)	1–5000
		5,001–20,000
		20,001–50,000
		50,001–100,000
		>100,000
7	Does your company have a data strategy?	Yes
		No
8	How much does your data strategy influence your company's strategic decisions?	Not at all
		Very weak
		Relatively weak
		Predominantly
		Relatively
		Relatively strong
9	How much does your data strategy influence your daily business?	Very strong
		Not at all
		Very weak
		Relatively weak
		Predominantly
10	How do you assess the added value of your data strategy about your business goals?	Relatively
		Relatively strong
		Very strong
		Very low (0–10%)
		Low (11–25%)
		Below average (26–40%)
11	How comprehensively are roles and responsibilities related to data implemented in your organization?	Average (41–60%)
		Above average (61–75%)
		High (76–90%)
		Very high (91–100%)
		Not at all
		Very weak
12	How comprehensively are data processes (e.g., creation of data, use of data) implemented in your company?	Relatively weak
		Predominantly
		Relatively
		Relatively strong
		Very strong

**Table A2.** *Cont.*

Nr.	Question	Possible Answers
13	How much is the handling of data influenced by company policies and standards?	Not at all Very weak Relatively weak Predominantly Relatively Relatively strong Very strong
14	How automated are your data processes (e.g., creation of new master data)?	Not at all Very weak Relatively weak Predominantly Relatively Relatively strong Very strong
15	To what extent does your company address the issue of data architecture?	Not at all Very weak Relatively weak Predominantly Relatively Relatively strong Very strong
16	How comprehensively are your metadata processes implemented?	Not at all Very weak Relatively weak Predominantly Relatively Relatively strong Very strong
17	Do you plan to implement a specific data architecture in your company, or have you already implemented a data architecture?	Yes No I don't know
18	If you answer yes, which one?	Open text answer
19	How often are decisions made in your organization based on opinions or gut feelings?	Never Hardly Rarely Occasionally Predominantly Mostly Always
20	How well are your employees trained in data and analytics skills?	Not at all Very weak Relatively weak Predominantly Relatively Relatively strong Very strong
21	How often are data and analytics used to make decisions in the company?	Never Hardly Rarely Occasionally Predominantly Mostly Always

Table A2. Cont.

Nr.	Question	Possible Answers
22	How often are decisions evaluated retrospectively based on data and analytics?	Never
		Hardly
		Rarely
		Occasionally
		Predominantly
		Mostly
		Always
23	Would you describe your company as data literate?	Yes
		No
24	How often do you use data to develop new business models?	Never
		Hardly
		Rarely
		Occasionally
		Predominantly
		Mostly
		Always
25	How regularly is data quality measured in your company?	Never
		Hardly
		Rarely
		Occasionally
		Predominantly
		Mostly
		Always
26	How would you rate the current data quality in your company?	Very low (0–10%)
		Low (11–25%)
		Below average (26–40%)
		Average (41–60%)
		Above average (61–75%)
		High (76–90%)
		Very high (91–100%)
27	How would you rate your company's business performance in terms of classical financial indicators (sales, revenue) over the past year?	Not at all
		Very weak
		Relatively weak
		Pre-dominantly
		Relatively
		Relatively strong
		Very strong
28	How do you assess your company's future competitiveness?	Very low (0–10%)
		Low (11–25%)
		Below average (26–40%)
		Average (41–60%)
		Above average (61–75%)
		High (76–90%)
		Very high (91–100%)
29	How do you evaluate the innovative power of your company?	Very low (0–10%)
		Low (11–25%)
		Below average (26–40%)
		Average (41–60%)
		Above average (61–75%)
		High (76–90%)
		Very high (91–100%)

**Table A2.** *Cont.*

Nr.	Question	Possible Answers
30	How do you rate the employee satisfaction in your company?	Very low (0–10%)
		Low (11–25%)
		Below average (26–40%)
		Average (41–60%)
		Above average (61–75%)
		High (76–90%)
		Very high (91–100%)
31	How well are you able to understand the needs and requirements of your customers and offer appropriate solutions?	Not at all
		Very weak
		Relatively weak
		Predominantly
		Relatively
		Relatively strong
		Very strong
32	How much has your company invested in digitalization in recent years?	Not at all
		Very weak
		Relatively weak
		Predominantly
		Relatively
		Relatively strong
		Very strong
33	How do they assess the degree of digitalization in your company?	Very low (0–10%)
		Low (11–25%)
		Below average (26–40%)
		Average (41–60%)
		Above average (61–75%)
		High (76–90%)
		Very high (91–100%)
34	If you have further comments or questions, please feel free to share them with us.	Open text answer

**Table A3.** Evaluation matrix based on Saaty’s scale.

D1—Strategy	Data Strategy and Roadmap	Business and IT Strategy Alignment	Cultural Fit of Strategy
Data Strategy and Roadmap	1	0.5	2
Business and IT Strategy Alignment	2	1	2
Cultural Fit of Strategy	0.5	0.5	1

**Table A4.** Geometric values of process dimension strategy.

D1—Strategy	Data Strategy and Roadmap	Business and IT Strategy Alignment	Cultural Fit of Strategy
Data Strategy and Roadmap	1	1.0699132	2.220906
Business and IT Strategy Alignment	0.9346553	1	1.414214
Cultural Fit of Strategy	0.4502667	0.7071068	1

**Table A5.** Normalized matrix.

D1—Strategy	Data Strategy and Roadmap	Business and IT Strategy Alignment	Cultural Fit of Strategy
Data Strategy and Roadmap	0.7182858 + 0i	0.7182858 + 0.0000000i	0.7182858 + 0.0000000i
Business and IT Strategy Alignment	0.5907357 + 0i	−0.2953678 − 0.5115921i	−0.2953678 + 0.5115921i
Cultural Fit of Strategy	0.3675553 + 0i	−0.1837776 + 0.3183122i	−0.1837776 − 0.3183122i

**Table A6.** Short data strategy assessment.

Nr.	Question	Possible Answers
1	Before we start your assessment, we want to know one detail: In which industry are you working?	Pharma
		Manufacturing
		Automotive
		Consumer products and goods
		Life sciences
		Energy
		Chemicals
2	Do you want a benchmark with some recommendation for action? Please let us know your mail address below.	Open text answer
3	How would you describe your data strategy?	1. Incomplete: no data strategy in place or planned
		2. Performed: initiative for establishing a data strategy is planned
		3. Managed: initiative for a data strategy roadmap started by assessing the current status quo
		4. Established: a holistic data strategy is defined and communicated—for example: a roadmap and defined working packages
		5. Predictable: the profound data strategy is integrated into the yearly overarching company strategy process
		6. Innovating: the data strategy is continuously updated regarding trends/innovations and has a lasting influence for the whole company in terms of value creation
4	How do you rate your organization's data management processes (data lifecycle, data maintenance, etc.)?	1. Incomplete: no established lifecycle or maintenance processes; no integration into an end-to-end approach
		2. Performed: processes exist, but are not documented/standardized
		3. Managed: processes exist, but only minimal end-to-end process integration
		4. Established: some established processes are partially integrated into the end-to-end approach
		5. Predictable: processes are very well established and almost fully integrated into the end-to-end approach
		6. Innovating: data processes are not only well defined but actively managed and refined as well as fully integrated into the end-to-end approach, resulting in efficient overall process execution
5	How mature is your data organization and data literacy skills?	1. Incomplete: no data roles/committees or anything else; no willingness to invest in data resources
		2. Performed: a first step into a data organization is planned; some distributed knowledge exists within company
		3. Managed: future data organization (roles, committees, incl. enablement and training) is conceptualized
		4. Established: data organization is implemented and communicated; most important stakeholders are enabled
		5. Predictable: an overarching data organization is in place, the needed enablement is done; a data culture is established
		6. Innovating: the established data organization is reviewed and adapted continuously; data competencies are spread all over; a data culture is well established and lived within the company
6	How advanced is your data standardization, harmonization, and normalization?	1. Incomplete: no approach
		2. Performed: data has been mapped across key systems
		3. Managed: initial successful attempts at harmonizing data across some systems
		4. Established: confirmed process to unify data has tackled main key data objects
		5. Predictable: advanced process has standardized and harmonized the majority of data across key systems
		6. Innovating: transparent cross-system data model including standards and guidelines at attribute level
7	Is there a standardized approach for continuous data quality management and data improvement introduced?	1. Incomplete: no standardized approach
		2. Performed: some occasional initiatives were/are started to improve data and its quality, but no systematic approach
		3. Managed: standardized data quality management process and data improvement process are planned and conceptualized
		4. Established: settled approach with some ongoing data quality management and improvement
		5. Predictable: data quality management and continuous data improvement are actively managed and standardized enterprise-wide
		6. Innovating: data quality management and continuous data improvement is used to gain a huge competitive advantage

**Table A6.** *Cont.*

Nr.	Question	Possible Answers
8	How managed is the integration among diverse systems (data context), and is there a structured interface in place?	<p>1. Incomplete: system integration/consolidation is infancy, with little coordination or comprehensive planning</p> <p>2. Performed: some initial efforts in system integration/consolidation</p> <p>3. Managed: system integration/consolidation is showing signs of thoughtful consideration but lacks full coordination</p> <p>4. Established: system integration/consolidation efforts are well thought out and coordinated, but there is room for improvement in terms of comprehensiveness</p> <p>5. Predictable: system integration/consolidation is well planned and coordinated and exhibits a comprehensive approach</p> <p>6. Innovating: system integration and interfaces are highly mature, with optimal planning, coordination, and a comprehensive approach to meet business goals</p>
9	How does your organization enable the value generation through innovation, automation, and AI/data use case management?	<p>1. Incomplete: no initiative that guides through innovation or a data value generation process</p> <p>2. Performed: innovation and automation or some AI use cases are in discussion anywhere in the organization, but only few people are informed</p> <p>3. Managed: somewhere within the company, processes for ideating and developing innovative data solutions exist but are not standardized</p> <p>4. Established: based on market research and/or (customer) needs, innovative data methods/solutions (e.g., AI, automation, etc.) are used within the whole company</p> <p>5. Predictable: a structured approach to identify value generating data methods/solutions (incl. use case management) is established organization-wide</p> <p>6. Innovating: highly structured process actively enables value generation through innovation, automation, and the ideation and development of data and AI use cases</p>
10	Which data topics would you rank the most important (and needed) within your company? Rank between 0 and 10 possible	<p>Data-driven decision making</p> <p>Data analytics and AI</p> <p>Master data management</p> <p>Data governance and organization</p> <p>Data strategy and roadmap</p> <p>Data quality</p> <p>Meta data management</p> <p>Data literacy and coaching</p> <p>Automation and workflows</p> <p>Reference data management</p>
11	What are the three biggest challenges in the context of data you are dealing with? Please name and rank them.	<p>Open text answer (rank nr. 1)</p> <p>Open text answer (rank nr. 2)</p> <p>Open text answer (rank nr. 3)</p>

**References**

- Provost, F.; Fawcett, T. Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data* **2013**, *1*, 51–59. [CrossRef] [PubMed]
- Kim, M. Software Engineering for Data Analytics. *IEEE Softw.* **2020**, *37*, 36–42. [CrossRef]
- Big Data, Analytics and the Path from Insights to—ProQuest. Available online: <https://www.proquest.com/docview/845235605?pq-origsite=gscholar&sourcetype=Scholarly%20Journals> (accessed on 18 February 2025).
- Hupperz, M.; Gür, I.; Möller, F.; Otto, B. What is a Data-Driven Organization? In Proceedings of the Twenty-Seventh Americas Conference on Information Systems, Online, 9–13 August 2021.
- Sarshar, M.; Finnemore, M.; Haigh, R.; Goulding, J. *Spice: Is a Capability Maturity Model Applicable in the Construction Industry?* Institute for Research in Construction: Ottawa ON, Canada, 1999; p. 9.
- Baolong, Y.; Hong, W.; Haodong, Z. Research and application of data management based on Data Management Maturity Model (DMM). In Proceedings of the 2018 10th International Conference on Machine Learning and Computing, Macau, China, 26–28 February 2018; ACM: Macau, China, 2018; pp. 157–160. [CrossRef]
- ISO/IEC 15504-5:2012; Information Technology—Process Assessment. IOP: Bristol, UK, 2012. Available online: <https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/05/60555.html> (accessed on 8 August 2024).
- Gökalp, M.O.; Kayabay, K.; Gökalp, E.; Koçyiğit, A.; Eren, P.E. Assessment of process capabilities in transition to a data-driven organisation: A multidisciplinary approach. *IET Softw.* **2021**, *15*, 376–390. [CrossRef]

9. Pörtner, L.; Möske, R.; Riel, A. Data Management Strategy Assessment for Leveraging the Digital Transformation: A Comparison Between Two Models: DX-CMM and Camelot DMM. In *Systems, Software and Services Process Improvement*; Yilmaz, M., Clarke, P., Messnarz, R., Wöran, B., Eds.; Springer International Publishing: Cham, Switzerland, 2022; pp. 553–567.
10. Luber, S. Was ist Data Management/Datenmanagement? Available online: <https://www.storage-insider.de/was-ist-data-managementdatenmanagement-a-850258/> (accessed on 1 September 2024).
11. Was ist Datenmanagement? | Definition, Bedeutung, Prozesse | SAP Insights. SAP. Available online: <https://www.sap.com/germany/products/technology-platform/what-is-data-management.html> (accessed on 18 February 2024).
12. Otto, B.; KHüner, M.; Österle, D.H. *Funktionsarchitektur für Unternehmensweites Stammdatenmanagement*; Institut für Wirtschaftsinformatik, Universität St. Gallen: St. Gallen, Switzerland, 2009; p. 69.
13. Scheuch, R.; Gansor, T.; Ziller, C. *Master Data Management: Strategie, Organisation, Architektur*; dpunkt.verlag: Heidelberg, Germany, 2012; ISBN 978-3-86491-165-1.
14. de Bruin, T.; Freeze, R.; Kulkarni, U.; Rosemann, M. Understanding the Main Phases of Developing a Maturity Assessment Model. In Proceedings of the ACIS 2005, Jeju Island, Republic of Korea, 14–16 July 2005.
15. Packowski, J. (Ed.) *Strategic Master Data Management: Prerequisite for Agile and Efficient Business Processes*; Study Findings; Camelot Management Consultants AG: Mannheim, Germany, 2012.
16. Azizan, S.; Ismail, R.; Baharum, A.; Hidayah Mat Zain, N. Exploring the Factors That Influence the Success of Digitalization in an Organization’s IT Department. In Proceedings of the 2021 6th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE), Kedah, Malaysia, 1–3 December 2021; Volume 6, pp. 1–6. [CrossRef]
17. International Assessor Certification Scheme e.V. (Intacs). *Data Management SPICE*; International Assessor Certification Scheme e.V. (Intacs): Köln, Germany, 2022.
18. DMMM: Data Management Maturity Model | IEEE Conference Publication | IEEE Xplore. Available online: <https://ieeexplore.ieee.org/document/9635003> (accessed on 8 August 2024).
19. Ko, C.; Adywiratama, A.D.; Hidayanto, A.N. Master Data Management Maturity Model (MD3M) Assessment: A Case Study in Secretariat of Presidential Advisory Council. In Proceedings of the 2021 9th International Conference on Information and Communication Technology (ICoICT), Yogyakarta, Indonesia, 3–5 August 2021; pp. 359–363. [CrossRef]
20. Kurniawan, D.H.; Ruldeviyani, Y.; Adrian, M.R.; Handayani, S.; Pohan, M.R.; Khairunnisa, T.R. Data Governance Maturity Assessment: A Case Study in IT Bureau of Audit Board. In Proceedings of the 2019 International Conference on Information Management and Technology (ICIMTech), Bali, Indonesia, 19–20 August 2019; Volume 1, pp. 629–634. [CrossRef]
21. Iqbal, R.; Yuda, P.; Aditya, W.; Hidayanto, A.N.; Wuri Handayani, P.; Harahap, N.C. Master Data Management Maturity Assessment: Case Study of XYZ Company. In Proceedings of the 2019 2nd International Conference on Applied Information Technology and Innovation (ICAITI), Bali, Indonesia, 21–22 September 2019; pp. 133–139. [CrossRef]
22. Pinto, M.R.; Salume, P.K.; Barbosa, M.W.; de Sousa, P.R. The path to digital maturity: A cluster analysis of the retail industry in an emerging economy. *Technol. Soc.* **2023**, *72*, 102191. [CrossRef]
23. Nyonawan, M.; Suharjo; Utama, D.N. Evaluation of Information Technology Governance in STMIK Mikroskil Using COBIT 5 Framework. In Proceedings of the 2018 International Conference on Information Management and Technology (ICIMTech), Jakarta, Indonesia, 3–5 September 2018; pp. 137–142. [CrossRef]
24. Berndtsson, M.; Forsberg, D.; Stein, D.; Svahn, T. Becoming a data-driven organisation. In Proceedings of the 26th European Conference on Information Systems: Beyond Digitization—Facets of Socio-Technical Change, Portsmouth, UK, 23–28 June 2018; Association for Information Systems: Atlanta, GA, USA, 2018.
25. *ISO/IEC 33002:2015; Information Technology—Process Assessment—Requirements for Performing Process Assessment*. VDE VERLAG: Berlin, Germany, 2015. Available online: <https://www.vde-verlag.de/iec-normen/221509/iso-iec-33002-2015.html> (accessed on 18 February 2025).
26. Mühlbacher, A.C.; Kaczynski, A. Der Analytic Hierarchy Process (AHP): Eine Methode zur Entscheidungsunterstützung im Gesundheitswesen. *PharmacoEcon. Ger. Res. Artic.* **2013**, *11*, 119–132. [CrossRef]
27. Peters, M.L.; Zelewski, S. *Analytical Hierarchy Process (AHP)—Dargestellt am Beispiel der Auswahl von Projektmanagement-Software zum Multiprojektmanagement; Arbeitsbericht Nr. 14*; University of Duisburg-Essen: Essen, Germany, 2022.
28. Golden, B.L.; Wang, Q. An Alternate Measure of Consistency. In *The Analytic Hierarchy Process: Applications and Studies*; Golden, B.L., Wasil, E.A., Harker, P.T., Eds.; Springer: Berlin/Heidelberg, Germany, 1989; pp. 68–81, ISBN 978-3-642-50244-6.
29. Belghith, O.; Skhiri, S.; Zitoun, S.; Ferjaoui, S. A Survey of Maturity Models in Data Management. In Proceedings of the 2021 IEEE 12th International Conference on Mechanical and Intelligent Manufacturing Technologies (ICMIMT), Cape Town, South Africa, 13–15 May 2021; pp. 298–309.
30. Smith, A. *Enterprise Information Management Maturity: Data Governance’s Role—Enterprise Information Management Institute*; EIM Institute: Valletta, Malta, 2009. Available online: <http://www.eiminstitute.org/library/eimi-archives/volume-3-issue-1-january-2009-edition/EIM-Maturity/> (accessed on 1 July 2024).

31. Data Strategy Professionals. Available online: <https://www.datastrategypros.com/resources/data-management-maturity-assessment-dmma> (accessed on 1 July 2024).
32. Data Management—DCAM—EDM Council. Available online: <https://edmcouncil.org/frameworks/dcam/> (accessed on 2 July 2024).
33. Spruit, M.; Pietzka, K. MD3M: The master data management maturity model. *Comput. Hum. Behav.* **2015**, *51*, 1068–1076. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.