A Survey of Maturity Models in Data Management

Oumaima Belghith
R&D. Eura Nova
Eura Nova
Tunis, Tunisia
oumaima.belghith@euranova.eu

Sabri Skhiri
R&D. Eura Nova
Eura Nova
Mont-Saint-Guilbert, Belgium
sabri.skhiri@euranova.eu

Sirine Zitoun
R&D. Eura Nova
Eura Nova
Tunis, Tunisia
sirine.zitoun@euranova.eu

Syrine Ferjaoui
R&D. Eura Nova
Eura Nova
Tunis, Tunisia
syrine.ferjaoui@euranova.eu

***Abstract*—Maturity models are helpful business tools that refine and develop the way organizations conduct their businesses and benchmark their maturity status against a scale or with industry peers. They serve to better prioritize the actions for improvement and control the progress in reaching the target maturity stage. To the best of our knowledge, very few survey papers are available on data management maturity models in academia, from which we studied their data and findings. In this context, our paper summarizes and organizes a variety of research that is related to or encompasses the data management field. Consequently, this paper is of interest both for scientists as well as practitioners from different industries and fields as it aims to highlight the importance of maturity models in the field of data management. From an academic perspective, our survey delivers a thorough literature review as it investigates maturity models that are either for or related to data management. Moreover, it offers a comparative analysis in terms of the main concepts and features associated with these models through a developed metamodel. This proposed framework describes the functional coverage of data management maturity models where models can be compared and evaluated based on their approaches to identify and categorize the data management related functions. As a result, this metamodel can serve as a tool for researchers who can exploit this framework to position future maturity models.**

Keywords—maturity models; data management; maturity assessment; digital transformation

I. INTRODUCTION

Enterprises are currently facing an increasingly competitive and erratic global business environment, aligning data projects at the edge of technical and organizational projects. They, consequently, encounter changes in IT organization as well as in the operating models and organizational structure.

Moreover, organizations struggle to identify and build the appropriate capabilities needed to manage and deploy the massive influx of data, due to being unaware of where they stand in terms of data management maturity.

Therefore, the assessment of the digital transformation's progress is an important step that every organization needs to understand and go through via the tracking of processes' improvements and the evaluation of the level of maturity in terms of growth and capabilities. And thus, to figure out how to evolve from their current state to become mature, gain a competitive advantage, and maximize their business value at the same time, while keeping the focus on the business needs and strategic goals, these organizations should benefit from maturity models.

There is extensive literature produced by various authors in the past decade as well as the development that led to the current thinking and practices of maturity models and their significance as assessment and improvement tools, for several business domains [1], [2]. However, for the data management field, according to what we reportedly believe, we focus and highlight in our study only the most relevant literature works which were conducted throughout the years, while offering new perspectives in terms of analysis and presentation on this area that requires additional investigation and research in the future.

The following section further clarifies the general context of our research from a business perspective and provides the key concepts and terms related to maturity models in the field of data management. A third section is dedicated to the research methodology adopted to analyze the studied models, using a set of assessment criteria, and in the elaboration of a metamodel that illustrates some of the conceptual similarities and differences between them. Afterward, the results of the survey are presented through an overview of a sample of models and then a discussion of the findings from the extensive analysis is presented. Lastly, a conclusion on the entire process is drawn and contributions are mentioned.

II. BACKGROUND

To maintain a competitive advantage, companies strive to provide the best experience for their customers and employees. This is achieved through the modernization and automation of processes and operations through digital platforms [3]. Thus, the key factor is data as described in this context as "The Gold" in this current era by Everest Group's CEO who shared that the global market for data and data analytics is expected to explode [3]. This was further affirmed by a research report on the enterprise data management market which was published by MarketsandMarkets [4]. The study showed that the size of this market will considerably witness significant growth from USD 77.9 billion in 2020 to USD 122.9 billion by 2025. Nowadays, organizations strive to sustain a certain level of advantage by adopting new technologies and integrating them in their transactional and analytical operations as well as their customer experience [4].

Consequently, companies need to monitor and benchmark their data analytics capabilities regularly, which requires the

measurement of their business value and comparison with similar systems in other companies. As data are useful only when put into context, another crucial key factor for success is matching the current maturity level within the organization to the appropriate metrics and pacing its growth via the integrated maturity model. Only then, the benefit of having a competitive advantage will be the highest [5]. Thus, it is most crucial for an organization to realize this correlation and be aware of the current maturity level of data management and of what needs to be achieved to reach the next level.

Therefore, employing the right tools to assess the level of data management maturity is a critical step in the digital transformation journey of most data-driven companies. In this context, maturity models serve as tools that define levels of efficiency, definition, manageability, and measurement of the monitored environment selected by the organization [6].

Having the assistance of such models facilitates the understanding of where a company stands in terms of evolving and grants it a competitive advantage through the offer of an adequate baseline for benchmarking and tracking the evolution of maturity assessment in comparison with industrial competitors or intercompany.

A. Literature Review

Data management maturity assessment is "a method for ranking practices for handling data within an organization to characterize the current state of data management and its impact on the organization" [7].

With that being a research trigger, and to further discuss maturity models, the next section will firstly give an outlook on data management and data governance, followed by the evolution of the capability maturity concept and maturity models.

1) Data Management: Organizations recognize their data as vital assets, a valuable tool that helps businesses bring not only innovative but also strategic goals. Despite its importance and recognition, data management is still deemed one of the most vital issues in the current technological landscape, and few businesses successfully derive value from it [8].

Previous research debated on the identification of the reasoning behind the failure of businesses to manage their information assets, hence several barriers were found to inhibit the issue. Some of them were associated more with the people and culture of an organization, such as lack of communication, knowledge-sharing initiatives, and proficiency and motivation [9]. While others, [10], debated that the barriers relate mostly to the incompetent organizational management and leadership, organization culture, oblivion towards costs, value, and benefits. However, to overcome all, regardless of their grounds, data, and information needed to be recognized as fundamental assets to day-to-day operations. Nowadays, they are the key assets for organizations to rely on in their decision-making process and operational processes. But for them to be potent assets, they need to be of good quality and well-managed [11].

It has been agreed that data management needs to encompass all major disciplines involved in making the data accessible, reliable, and appropriate for internal as well as external use. Nevertheless, it is ought to balance the strategic needs of an enterprise with its operational ones. Thus, as mentioned in DAMA International [7], the best way to achieve this balance is by embracing dedication towards the successful and effective execution of the management of data.

2) Data Governance: The term 'Governance' refers to the process with which an organization strives to ensure that its strategies are set, monitored, and achieved [12].

The definition of data governance differed across researchers. However, the Data Governance Institute (DGI) defined it in the following way:

"*Data governance is a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with* what *information, and when, under what circumstances, using what methods*" [13].

New technologies have enabled the collection and use of Big Data, and the importance of data ethics has grown along with humans' ability to exploit the vast amount of data and information produced as part of our daily life. As a result, data governance has been a structure in many organizations. The objective behind it is to achieve a proper and sustainable management process of data as an asset, following policies, and best practices [14]. In addition to that, the implementation of a good data governance framework would enable any organization to manage and align its data-related processes with the organizational vision and mission, strategy, values, and culture [15].

3) Maturity Models*:* The concept of maturity is based on the Capability Maturity Model (CMM) which was developed in 1984 by the Software Engineering Institute (SEI) to provide an innovative methodology and tools for improving and establishing software development processes [16].

Consequently, maturity became a term designed for the process conducted in the measurement of capabilities of an organization in a particular business area. It is the collection of functions and processes by which organizations can accurately identify, define, integrate, protect, effectively collect, and eventually deliver data that is suitable and fit for internal applications as well as external usage [17].

The success of the pioneer CMMI [18] has been an important inspiration, resulting in the literature research field that has revealed its rising academic interest, as well as the increasing overall adoption of models in practice. This substantially remarkable interest in the domain thereby, explains the significant number of developed and proposed models by various software companies and consultancies across multiple diverse domains.

For decades, organizations are rooted in the high importance of digital transformation orientation and continuous data integration in organizational design. However, they failed to allocate the impacts of such deployments and improvements. The constraint of the intractability of their processes' progress induces the importance of the model's implementation and its evaluation's assessment.

As a result, model developers have been striving to re-establish the lack of key tools when developing and deploying their models, to help organizations assess their level of maturity [19].

According to the data management practitioner, Irina Steenbeek [20], by implementing a maturity assessment, a company will be able to identify its "As-Is" status, create a foundation for the development of strategies, roadmaps, plans, and actions to reach the 'To Be' status as well as achieve impactful and efficient processes. A data management maturity model is an effective tool to measure the capabilities' maturity of an organization in a particular business or technical area. Besides, different authors [21] have shared the same identified advantages that maturity models' integration results are the enabler of a pathway towards continuous improvement and effective growth, by being a facilitator in terms of where an organization should put its first efforts and allocate its capabilities along with resources. Furthermore, Steenbeek [20] emphasized that amongst other things, models are recognized for sharing a common development feature which is defining a set of dimensions, most likely a collection of major key process areas that encompass minor functions or attributes.

Additionally, studies throughout the years proved that most maturity levels for the assessment of an organization's position are CMM-Based models [19]. Regarding the maturity evaluation methodology and assessment criteria, almost all models have the same scoring system with 5 incremented stages, labeled levels. The rate of evaluation is usually a five-stage scale from 0 or 1 being the lowest, to 5 or 6 being the highest. Labeled and explained as such, from lowest to highest.

III. RESEARCH METHODOLOGY

This review provides a critical and more in-depth analysis of the extensive literature produced and developed that led to the current practices in data and information management and governance fields. The conducted literature review highlighted new perspectives on data management maturity models which is an area that requires ongoing investigation and research.

To gather the relevant information, we complemented our research by exploring a set of versatile initially established works, from academia as well as the industry, from the first introduced model in 1986 until today, to further figure out the methodology and development behind data management maturity models.

Throughout this process, the selection approach was based on a set of criteria including the identification of different maturity models used to assess organizations in various business areas, planning for improvement during their digital transformation as well as clarifying the internal data management procedures and practices used.

A. Assessment Criteria

The analysis[1] was based on a set of criteria. In total, 20 models have met our selection criteria. We used them to form the metamodel that served as a highlight to their commonalities and differences regarding focus domains, capabilities, and sub-capabilities after analyzing to compare their characteristics.

The analysis is divided into four main parts, to enable assessing and then comparing the selected maturity models while providing details on some of the strengths and weaknesses associated with each model.

These attributes serve as guidelines by which maturity models can be classified and examined. Namely, the models' structure, assessment processes, their outputs, and means of implementation, support details, and more general features. These different attributes, assigned to each model, are always taken into consideration by organizations, who have already formalized the data management function or those who have not yet. Each aspect serves as an important differentiator for them to select the most suitable model that fits well with their vision and focus and aligns best with their business needs.

Examples of the models' assessment are provided with each set of criteria.

1) Model Structure: The first section focuses mainly on the structural aspect of the models as it contains:

   a) The **origin** and the **main references**

   b) The **number** and **names** of the **attributes** and **maturity levels** details

   c) The **practicality** of each model that we defined regarding whether it is **general or problem specific**.

   d) The **maturity definition:** as it differs from one model to the other.

2) Model Assessment Details: This second section focuses on the application methods provided by the models.

   e) Availability and description of the model's assessment process

   f) Identification of strong and weak points for an organization

   g) Provision of priority of improvement for the organizations.

   h) Assessment cost: free

   i) Assessment continuity: If the maturity model aims for a continuous assessment.

   j) Maturity levels' calculation/measuring details.

3) General Features: The features section is a description of how the models are predefined and established.

   a) The main goal(s) models enable organizations to achieve.
   b) Models framework description
   c) Tools that each model is integrated with.

4) Strengths and Weaknesses: This section of the analysis highlights the positive and negative points that each model is identified with.

[1] https://drive.google.com/file/d/1fF06NcBUw16sswATD9to4WZ3u8G2QvfE/view

These aspects will enable readers to swiftly compare between models and rapidly identify which is best for the assessment they wish to conduct in their organization.

B. Analysis Approach

A systematic comparison between the studied maturity models, based on specific criteria, has led to a clearer understanding of how these models are structured in terms of areas of capability, assessment methods, and features. The purpose of this comparison is to evaluate the different approaches and use them as input for the development of the company's maturity model.

The metamodel, illustrated in Fig. 1, has been developed to compare the different maturity models that have been selected for analysis and their features.

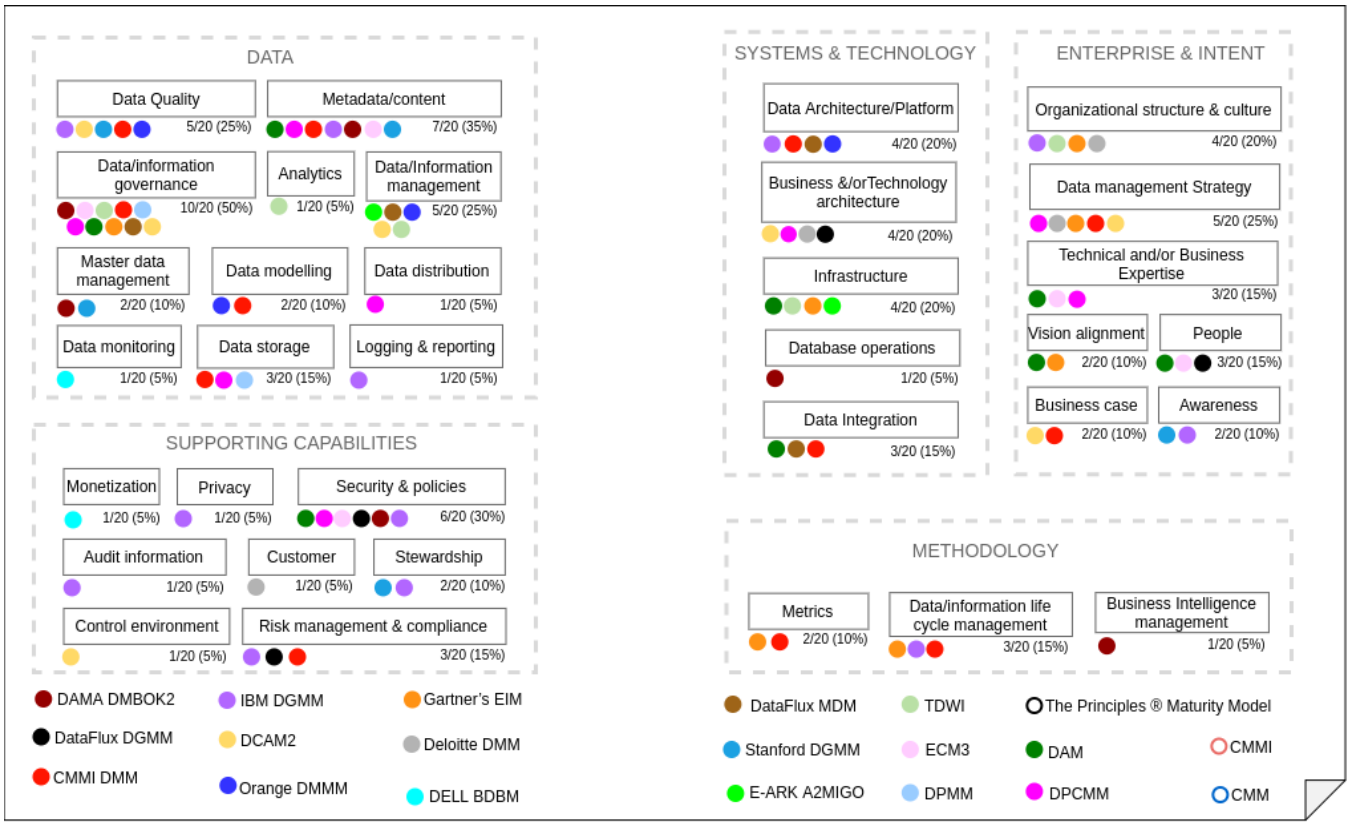

Fig. 1. Maturity models in data management metamodel.

With five main components to underline the chosen maturity models' focus areas, each component is a representation of the set of attributes that were assigned according to the related characteristics and capabilities established for each model. The "Systems & Technology" component underlines the features that relate to the technical tools and operations. The "Data" component represents the data and information related functions as well as their associated operations for management. The "Methodology" component underscores the technical methodical capabilities, as for the "Enterprise & Intent" component, it delineates the business, goals, and culture related functions. And finally, "Supporting Capabilities" which is the component that designates the supplementary features that some models established to support their main categories for an additional contribution.

Hence, through this approach, the metamodel illustrates similarities and commonalities, and how each model is still connected with another by sharing their practicality, feature, or focus domain, despite some of them having differently structured attributes and functions.

As depicted in our metamodel, some conceptual differences can be spotted between the models when it comes to interpretations and perspectives, along with their percentage.

A prime example of this is the variation in defining functions such as "data governance" as described by DCAM V2 [22], [23] as a "Business Capability", while CMMI's [24] DMM labels it as a "Category". Although some of the data management functions' roles are similar in some models, others are entirely different. This becomes obvious when considering the function "Integration" which is regarded as a data management "Process Area", which operates in the "Architecture" category in the CMMI's DMM [25]. While it is separated from "Architecture" by 15% of our selection, such as the DataFlux MDM [26], and is treated as an independent data management "Component". Some models also differ in recognizing which functions are deliverables of other functions, which is also selected by 20% of the models.

Another example is Deloitte's DMM [27], [28] and Gartner's EIM [29]-[31], even though they share the same domains, they only have one common attribute, which is

Organizational structure and culture capability in the Enterprise and Intent domain.

The two models differ in all others, such as relating their system and technology capabilities separately, as Deloitte's DMM specifies in Business and/or Technology Architecture, while EIM considers the latter as Infrastructure.

The model DAMA DMBOK2 [7] also, alters from the other models as it does not have any practicality in the Enterprise and Intent domain, nor the supporting capabilities. However, it contributes by focusing on attributes that none of the depicted models have, such as Data Operations in the Systems and Technology area, and Business intelligence management in the Methodology function.

A further case in point is DELL BDBM [32], which distinguishes itself from the other models by adopting different approaches to the different domains. The model instigates the monetization function as a support capability, with introducing Data monitoring as a separate attribute in the Data domain, unlike the majority of the other models, 50% of our selection, define this latter as a subordinate in the prime capabilities, Data/Information management, or Data/Information Governance.

Therefore, this metamodel can serve as the starting point for building a new data management maturity model. By regrouping most of the capabilities and sub-capabilities used by the depicted models, with five key components for data management having been determined. The components can establish the basis for deciding on the number and labels of the capabilities, sub-capabilities, and criteria used as well.

IV. RESULTS

A. *Overview of Maturity Models in Data Management available in the Literature*

In this part, we analyze the selected maturity models by grouping them into families according to their practicality and assessment focus, as illustrated in Appendix I.

The following is a brief introduction to each model.

- Capability Maturity Model

The CMM model was developed by the Software Engineering Institute of Carnegie-Mellon University. It represents a path of improvements recommended for software organizations that want to increase their software process capability as it identifies and prioritizes findings.

The model consists of 5 levels of maturity with key practices as to how most effectively implement and institutionalize the key process areas.

- DAMA-DMBOK2 Data Management Maturity Model

This is the most well-known and frequently used guide for data management as DAMA-DMBOK defines data management as a set of Knowledge Areas and uses the term 'business function' to specify their content. It combines knowledge of different data management related areas and especially the best practices and techniques.

- Capability Maturity Model Integration

The CMMI (Capability Maturity Model Institute) has developed its own CMMI which provides assessment criteria for the following data management areas:

- Data Management Strategy
- Data Governance
- Data Quality
- Platform and Architecture
- Data Operations
- Supporting Process

Within each of these dimensions, the model provides a set of sub-processes for evaluation.

- Data Management Capability Assessment Model V2

The Enterprise Data Management Council first developed the first version of their DCAM in 2014. It is an industry-standard, based on the practical experiences of the world's leading organizations. It consists of best operational practices in data management, with an understanding of business value, combined with the reality of operational implementation. It initially had 37 capabilities with 115 sub-capabilities.

The scoring system they assigned was based on the stakeholders' engagement, the formality of processes, and the existence of artifacts that demonstrate the achievement of capabilities.

Later, in 2019, they released a DCAM v2. The updated version removed duplications and streamlined the sub-dimensions across 31 capabilities, now organized into 7 components with offering a digital version of the model online. With a change in its framework that now encompasses Artificial Intelligence as well as Data ethics and Business Process Design.

- IBM Data Governance Council Maturity Model

DGMM was first introduced in 2007 by IBM. It is based on the inputs of 55 council organizations. Its main purpose is to help organizations build consistency and quality control in governance through proven business technologies, best practices, and collaborative methods within each of the 11 data governance domains. These domains are individually assessed and placed within their maturity level. This is convenient for organizations that initiate their digital transformation by starting at the level and dimension best suited for their business needs.

IBM Industry models: Developed by IBM, these industry models are a set of technical and business data models pre-designed to meet the needs in particular industries such as Banking and Financial Services, Insurance, Telecommunications, Retail, Energy and Utility, and Healthcare. These models enhance and accelerate the digital transformation and analytics experience by providing best practices and using pre-designed industry-specific content that helps in managing enterprise data existing in data warehouses or the data lake.

Some of the common features shared amongst these models include time to value, business agility, high-quality business data, and audit readiness. Examples of IBM Industry models:

- IBM Banking and Financial Markets Data Warehouse

- IBM Banking Process & Service Models
- IBM Insurance Information Warehouse
- IBM Data Model for Energy and Utilities
- IBM Unified Data Model for Healthcare

- Gartner's Enterprise Information Management Maturity Model

Introduced in 2008, Gartner had emphasized that this model is not a single project but a program that evolves. It was developed to provide organizations with guidance and an enterprise-wide approach to manage their information assets.

The EIM Maturity Model has 5 stages of maturity and 7 Dimensions: vision, strategy, metrics, information governance, organizations and roles, information lifecycle, enabling infrastructure with a set of criteria for each dimension in the form of questions.

- Data Management Maturity Model

CMMI Institute developed its own practitioner-based Model DMMM that has a well-defined framework of data management practices with six key categories. DMMM is also used by organizations to assess their current state and build customized roadmaps to improve their data management capability.

- The Digital Asset Management Maturity Model

The DAM model helps organizations identify their current state, where it is supposed to be, and how to reach the target state after performing a gap analysis to develop the evolution path to be followed. A description of how to do a self-assessment is provided through 3 steps that consist of the inventory of all the stakeholders, identifying internal champions, creating and administering questionnaires, getting answers to identify weaknesses, and finally, actions needed to be taken.

- Digital Preservation Capability Maturity Model

Preservica's DPCMM is based on capabilities. It is a flexible approach that can be adapted to an organization's requirements and resources. It Identifies the main requirements for digital continuity regarding the desired future target and the level of risk the organization is willing to take on and focus on digital preservation. The model has 5 incremental stages for 15 components.

- Digital Preservation Maturity Model

A Preservica's Maturity Model as well. This model however serves as an assessment tool for organizations to understand the difference between digital preservation and digital archiving. It aims to assess how vulnerable the organization is to digital information and highlights the risks behind it.

The Model has 3 main sections, each divided into layers representing their level.

- Big Data Business Model Maturity Index

DELL's BDBM index focuses on the business model rather than the technology model. It provides guidelines for organizations to measure how well they adopted Big Data as well as practical recommendations for the business.

The model focuses on 5 Business Phases. It is also presented as an animation that ties all the points together into a compelling, provocative, and visually engaging story.

- Deloitte Digital Maturity Model

The first industry-standard digital maturity assessment model developed in collaboration with the TM Forum with key contributions from industry and experts. DMM enables business leaders to assess where they are in their transformation journey, prioritize, and focus on the transformation actions and make impactful project investments. It is also the first model to benchmark against competitors and measure progress.

- Transforming Data with Intelligence

TDWI offered 5 different sub-models and assessments for organizations to assess their transformation phases.

- Analytics Maturity Model and Assessment Tool
- IoT data readiness Assessment
- Self-service analytics Maturity Model
- Advanced analytics Maturity Model
- Hadoop readiness assessment

- ECM Maturity Model

ECM3 is a tool for evaluating an organization's performance against the ECM capabilities' framework.

With 13 dimensions across 3 categories, the model provides roadmaps for optimizing performance and maturing organizational capabilities.

- E-ARK Information Governance Maturity Model

This maturity model consists of five maturity levels traced with the ISO 16363 Criterion with three attributes defined which are called dimensions. It provides a self-assessment questionnaire for each dimension to calculate maturity levels with each level having multiple sub-levels, but they differ from one level to another and from one dimension to the other. Each questionnaire provides an introduction to its purpose and how it will be analyzed with the possible answers.

- The Principles Maturity Model

ARMA International's Principles MM describes for each Principle the characteristics of effective records management at five distinct levels of development: substandard, in development, essential, proactive, transformational.

It has a set of principles that focus on 8 dimensions representing an organization's state and improvement in the transformation process, with 5 levels that evaluate the position the organization stands at.

- DataFlux Master Data Model MDM

The MDM is considered an evolutionary program. It is a whole process that involves understanding the essential capabilities necessary for a successful Master model deployment, and identification to the degree of maturity of those capabilities necessary to make the MDM actionable.

It provides a conceptual outline of its technical components, their various layers in terms of their maturity, for businesses to target their desired level of maturity, develop and build a

roadmap for its implementation that articulates the steps needed to take when assembling the program.

- DataFlux Data Governance Maturity Model

The Data governance Maturity Model was first developed by DataFlux in 2007 and has been since then adjusted and advanced to empathize a business perspective that is the driving force for the need for managing data as an enterprise asset, and the means such as organization, process, and technology to reach the desired levels of data quality. One of the strengths of this model is that it offers detailed descriptions like the fact that each maturity level has an associated profile detailing the four dimensions.

These descriptions are self-evident in terms of how to move up to the next level of maturity.

- The "Orange" Data Management Maturity Model

The analogy that orange is a hybrid between a pomelo and a mandarin inspired the name of the model "as it perfectly symbolized the attempt to cross the "pomelos" of data management metamodels with the "mandarins" of data management maturity metamodels."

This standard metamodel was developed for data management. implementation and maturity assessment. It is based on the similarities between some of the key industry models and 6 key principles. But most importantly, it is based on the concept of business capabilities specified by 4 components: data/ information, process, tools, and roles. Data management is treated as a separate business function that operates in close collaboration with IT and other business functions, with the focus on both internal and external groups of customers. The "orange" model can be applied in companies for the development of a data management function and a maturity scan.

- Data Management Family

The maturity models of this group share data management as their assessment focus. Even though their specific practicalities may differ, they all serve as tools that evaluate the maturity of data assets in organizations and how it can be developed. All models have 5 maturity levels except for DCAMv2 [23][29] which differs by having an additional level "Conceptual". For the "Orange" model, for instance, its development encompasses the analysis of all the aforementioned models, from which a metamodel has been derived to generate the most significant and efficient aspects that a maturity model must include.

TABLE I. COMMONALITIES OF DATA MANAGEMENT FAMILY

Levels	Ad-hoc/Initial - In development - Defined
Attributes	Data Management - Data Governance - Data Quality - Data Architecture - Processes
Approach	- Identification of strong & weak points - Assessment Support - Assessment continuity with roadmap development strategies
Strengths	- Flexibility & adaptability to company specifications - Well-defined & enriched frameworks - Guidance and details on features - Best practices, actions & recommendations for maturity level evolution
Weaknesses	N/A

- Data/Information Governance Family

This group includes models that are particularly focused on the governance of data and information assets. IBM's maturity model [33][34][35] was one of the pioneers to introduce and develop such a model with this particular focus. It is also the only model to have 11 attributes that are individually assessed, of which there are exceptionally risk management, privacy, and audit information functions. Moreover, it is the sole model with this practicality that does not offer training with its assessment nor cover the people and culture sides of organizations.

TABLE II. COMMONALITIES OF DATA/INFORMATION GOVERNANCE FAMILY

Levels	Reactive - Proactive - Managed - Optimized
Attributes	Information Governance - Data Quality - Processes - Policies
Approach	- Strong & weak points identified - Inputs diversity
Strengths	- Current State assessment - Scope definition based on priorities - Framework focus on Information Governance - Risks & resources allocation
Weaknesses	- No workshops

- Software Development Family

This family group is composed of two of the most known and referred to models specialized in software development. CMM [16] represents a path of improvements recommended for software organizations that aim to increase their software process capability as it identifies and prioritizes findings. The CMMI institute also developed the successor of the CMM model, CMMI [16], which provides assessment criteria for the following data management areas Data Management Strategy, Data Governance, Data Quality, Platform, and Architecture, Data Operations, Supporting Process.

TABLE III. COMMONALITIES OF SOFTWARE DEVELOPMENT FAMILY

Levels	Initial - Defined - Managed
Attributes	Process Areas
Approach	- Assessment continuity - Assessment support & training - Strong & weak points identification
Strengths	- Recommendations for software development & data integration - Guidance for maturity improvement

	- Accepted & global best practice for the management and delivery of quality software processes
Weaknesses	- Unavailable scale metrics - Require high resources and knowledge - No measurement procedures

- Digital Assessment Family

The maturity models of this family are dedicated to the management of digital assets, and how the preservation of digital continuity should be practiced within companies. This group's assessment tools share the fact that they offer a free of charge evaluation adherents to their commonly collective five maturity levels. Deloitte's DMM [27][28] however is the only charged assessment, as it offers organizations an intricate benchmark against their competitors as well as effective gaps' identifications in their digital progress with established key recommendations for evolution.

TABLE IV. COMMONALITIES OF DIGITAL ASSESSMENT FAMILY

Levels	N/A
Attributes	Strategy - Technology
Approach	- Strong & weak points identification - Assessment support
Strengths	- Current state assessment for digital assets - Capabilities gaps' analysis and identification - Guidance for capabilities' improvement
Weaknesses	- No process continuity

- Analytics Family

This family assembles the 5 tools developed by TDWI company [38], that assess and guide organizations in their analytical programs. However, the drawback of this group is that all the assessments are shared with all their users, which may result in the use of such information by competitors for beneficial practices and weakness awareness.

TABLE V. COMMONALITIES OF ANALYTICS FAMILY

Levels	Nascent - Pre-Adoption - Early Adoption - Corporate Adoption - Mature Visionary
Attributes	Organization - Infrastructure - Data management - Analytics - Governance
Approach	- Described assessment methodology - Available tools for interpretation of guidelines and assessments
Strengths	- Guidelines for all phases - Provided recommendations on future actions - Opportunity to compare results with other organizations and filter companies according to size or industry
Weaknesses	- No identification of strong & weak points - Unavailable scale metrics - No training supports - The available information on other companies can be accessed used by competitors - Limited framework focuses

- Business Performance Family

Maturity models of this group are focused on the business aspects of organizations. They are assessments for organizational performances that offer optimization and strategies for alignment of capabilities as progress guidelines. However, none of this family's members share the same maturity levels or assessment attributes as the others. But their practicality is common as they all focus on business processes and business initiatives.

TABLE VI. COMMONALITIES OF BUSINESS PERFORMANCE FAMILY

Levels	N/A
Attributes	N/A
Approach	- Identification of strong & weak points - Assessment continuity - Framework focus on business aspects and processes
Strengths	- Improvement & progress tools - Guidelines for phases, processes & business initiatives - Opportunity to compare results with other organizations
Weaknesses	- The available information on other firms can be accessed used by competitors

V. DISCUSSION

The analysis conducted has highlighted a great awareness of the different maturity models whose number has been increasing throughout the years. Despite sharing the same domain, that is data management, the models analyzed differ in terms of origins, tools of assessment, and assessment focus.

As emphasized in previous studies as well as concluded now, the most well-known data management or governance maturity models are DAMA-DMBOK [7], DCAMv2 [23], [24], IBM data governance council maturity model [33]-[35], Stanford's DGMM [37], Gartner's enterprise information management [33]-[35]. These models constitute the foundation of almost every developed new model. Though Models like DAMA-DMBOK [7], DCAMv2 [23], [24], DMM [25], The "orange" model [20], and CMMI [16] are applied to the field of data management, others like IBM [33]-[35], Gartner (EIM) [29]-[31], DataFlux DGMM [26], A2MIGO [42], DGMM [37], and Stanford [37] focus on information governance aspects.

There is never one model that fits all when it comes to maturity assessment. However, each model has strengths, can bring valuable perspectives, and serves as a foundation for planning the implementation of data management. Prime examples of this are the Scientific Data Management maturity model Stanford's DGMM [37], and IBM's DGMM [33]-[35], which were based on the Software Engineering Institute's (SEI) Capability Maturity Model (CMM) [16], [24], [44]. DPCMM [43] and E-ARK [39] models also drew from the ISO when constructing their respective dimensions and levels. This is also the case for A2MIGO [42] and "Orange" [20] models that

clearly show the use of CMMI [16] in their work and development.

Another determinant for such matters is the number and name of the maturity levels. For this feature, it has been concluded that the number of levels varies between 5 and 6. Another striking difference comes with names, types, and numbers of attributes. For instance, DAMA-DMBOK [7] operates via key practice areas or knowledge areas, DCAMv2 [23], [24] adheres to a business capability concept, Stanford [37] and CMMI [16] use processes as a basis, and IBM addresses competency whereas DELL [32] developed its domains as business phases. As for the rest of the models, they mostly use dimensions. The next features are the assessment method of the models, the means of its implementation as well as the supporting system provided for the integration.

Additionally, another feature is the level of adaptability in which the CMMI [16] model has been used successfully in master data management models as well as data governance, due to easy adaptability and integration. Moreover, regardless of their costs, some models offer a great service for their clients with their easy way of implementation and use as well as the support and training included. Concerning maturity levels, a common concept among the studied models is the progressive maturing from reactive to proactive or predictive. The latter is the long term sought out objective which consists of managing risk and anticipating and preparing for opportunities and threats. It also includes the preparation and development of strategies for future implementations to hedge against any type of threat or error. This finally renders the enterprise more agile, adaptive, and dynamic.

Looking at the dimensions, the models' approaches are different when it comes to determining and defining the key attributes, and despite the distinct naming of these attributes.

We find that few models share the same semantic labeling process as they are mainly assigned as data quality, data/information governance, information/data security, strategy, architecture, infrastructure, metadata, and technology. It is worth noting that the approaches to defining these dimensions are rather different. These attributes' association is either "domains" or "categories" in some models but are often used as "dimensions" or "deliverables" for these dimensions in other different models. For instance, "data management strategy", which appears in five models, is a deliverable of one of the main categories in CMMI's DMM [25]. However, it is considered as a dimension in Deloitte DMM [27], [28] and Gartner's EIM [29]-[31].

The comparison between the models is challenging as they have substantial differences in their metamodels; with each having its specific characteristics and factors. Hence, the family grouping according to models' functionalities that we chose to follow. Accordingly, while it may seem that models belonging to the same family with the same practicality may have common attributes and maturity levels' metrics and assessment methodologies, such as data management family, information/data governance management family, and analytics family, some groups are different with no commonalities like digital assessment family and business performance family. However, for the most, they seem to share an approach to their focus with organizations, which induces their similarities in the offered model's strengths as well as weaknesses. These aspects were greatly noticed in all the family groups.

Nonetheless, the distribution of our selected maturity models was more highlighted in the metamodel framework that we developed, as it shows that despite their different structures and functions, the models are similar in different aspects such as the functions' definition and the supporting capabilities chosen to include in their assessment methodology. The most present supporting capability in a high number of maturity models is Security & Policies. In addition to that, the selected maturity models tend to always include data/information governance as the main attribute to their framework as it is an important feature in their assessment when it comes to data management and related domains.

The metamodel has also illustrated major differences between the models regarding the interpretations and definitions of data management functions, their roles, and the manner with which they intersect. The selected models' division according to families confirmed such differences, as many models do not define their function the same way as their family members despite belonging to the same group of practicality. The importance and roles of their features also differ, as for some models, an attribute could be the "parent" [44], "key process area" [7], or "main category" [25] however for another, it is a "sub-capability" [20] or "dimension" [26].

VI. CONCLUSION

The comprehensive literature review has put more emphasis on the importance of maturity models in determining the current or desirable maturity state and how it proved to be a valuable asset and "a path towards an increasingly organized and systematic way of doing business" [1]. Although there is a lot of research done on maturity models in the fields of data and information management as well as governance, the academic literature focuses mostly on the scope and general features of the models rather than the supporting scorecard or assessment tool, and how maturity levels can be measured and identified.

In this context, this survey paper is a contribution to the research community in the field of data management and maturity models as it highlights the valuable combination of academic research and practice into the service of business needs.

Through this study, which has started with the existing knowledge from literature, we have sought to examine a large sample of diverse maturity models, both academic and industrial, that are either for or related to data/information management as well as governance. The main tool used for this analysis is the elaborated metamodel which allows the comparison between the studied models and their grouping into 5 main areas of focus classification under 6 families according to their practicality and assessment focus.

This consequently led to a deep discussion of findings. The results suggest that companies should be more attentive and direct more focus on data management principles and practices.

This journey made us aware of the significance of maturity models for organizations, regardless of their industries, as this

tool serves as guidance in their digital transformation and their data management performance.

For further research, we proposed taking a different approach with regards to picking the maturity models fit for the analysis. Maturity models that cover more areas other than data management, could offer a new perspective on how to manage and govern other aspects that are not necessarily data related but critical within any organization. Such aspects can relate to people and culture, business strategies, and processes.

VII. Open Challenges

Implementing a maturity model in a company is bound to bring changes in its culture, operation model, organizational structure, and most importantly, the way data is foreseen and manipulated. As a result, the evaluation conducted needs to encompass the traceability of data within companies and the steps needed for its progress. With that being said, developing a maturity model in the data management field necessitates expertise and knowledge in its different features and how to use them for the most suitable future state for each company and their capacities. Thus, it introduces more open challenges to researchers and developers, which requires an ongoing methodological process that combines continuous and futuristic glance at its implementation in different aspects of business fields. Therefore, it is necessary to dig into the rationale and frameworks conducted behind different models and their development processes. And through that, we understood that a maturity model is not a simple assessment tool, but a solution for organizations to assess their present capabilities and build strategies and factual roadmaps for improvement upon them while preserving the alignment with their main goals. Which has advocated us to develop our data management maturity model.

VIII. Outlook

The conducted literature review on this domain and the analysis of the numerous maturity models related to data management enabled us to understand further the major factor for such a tool and how it must focus and embed an enriched architecture. Thereby, this motivated us to develop methods and techniques for our maturity assessment model in the data management field, that will not only focus on the implementation of data but also the organizational intents such as strategies and cultures, as well as the operating systems and their deployment. Thus, our tool will allow organizations that either plan on embarking on their digital transformation journey or already have, to get a refined and deepened scope on a high number of areas.

Our model[2] encompasses four major categories that each includes the related capabilities and enablers, which are 14 in total and are also expanded into sub-capabilities with explanatory of their aspects & features to assess for the organization and at the same time, maintaining a grid of dependencies between them. For a further definition of our chosen categories, the first being Enterprise and Intent; This category revolves around the organization's set goals and the process of achieving them while progressing in the digital transformation journey and forging ahead with technology implementation. It also comprises the different cultural changes this journey brings with it, and how finally achieving a data-driven culture. Second is the Data Management category which induces the fact that Data and information are fundamental assets that every organization should constantly manage and monitor. It includes the different aspects that Data needs to derive the most efficient value from it, while aligned with the business goals. The third category is Systems which highlights the significance of data in the implemented tools and their operations while preserving conservative standardized functions. And lastly, the Data Operations category includes the methodologies behind the integrated processes across the organization, as well as their analytical deployment and contribution to the overall business process.

Additionally, as commonly referenced in our selected maturity models' analysis, the chosen evaluation system follows a five maturity levels system, with a score range equivalence for each of the levels, to result from the initial assessment and evaluate organizations on their current state of maturity.

Our methodology was based on the approach highlighted in the developed metamodel framework for the analysis of maturity models, which made it possible for us to determine the best practices and approaches to follow in our development process. The conceptual differences and similarities resulted from the tool served as a guideline in the perception and recognition of what a maturity model in this field must include as well as offer. It also facilitated the selection of the significant features to escort our established grid as well as the linkages between them.

Acknowledgment

This work took part in a larger research & development effort planned by EURA NOVA to develop and introduce its own data management maturity model. The entire process entailed doing research through state-of-the-art and investigating the existing models in this domain.

References

[1] Samuel, P. B. (2019, November 26). Data Analytics and Data Management Market. Retrieved from https://www.forbes.com/sites/peterbendorsamuel/2019/11/26/dataanalytics-and-data-management-market/#1a0f36e77678.

[2] NC., M. (n.d.). PRESS RELEASES: Enterprise Data Management Market worth $122.9 billion by 2025. Retrieved from MarketsandMarkets: https://www.marketsandmarkets.com/PressReleases/enterprise-data-management.asp

[3] Cermak, T., Dochtermann, D., & Jesus-Olhausen, A. L. (2011). Introduction to a project portfolio management maturity model. Paper presented at PMI® Global Congress 2011 —North America, Dallas, TX. Newtown Square, PA: Project Management Institute

[4] Rajterič, I. H. (2010). OVERVIEW OF BUSINESS INTELLIGENCE MATURITY. Management, 15, 47-67. Retrieved from https://hrcak.srce.hr/file/81745.

[5] DAMA International. (2017). DAMA-DMBOK2 Data Management Body of Knowledge (2nd ed.). (S. E.-C. Deborah Henderson, Ed.) NJ, USA: Technics Publications. Retrieved 2009

[2] https://docs.google.com/document/d/1IqFh7WuI4OLLgSoeTbI9vI6iDUB_BwPv60EXoPlm5Yc/edit?usp=sharing

[6] Luftman, J., Derksen, B., Dwivedi, R., Santana, M., Zadeh, H., & Rigoni, E. (2015, September 1). Influential IT Management Trends: An International Study. Journal of Information Technology: Sage Journals, 30, 293-305. Retrieved from journals.sagepub.com.

[7] Oliver, G. (2014). Organizational Culture for Information Managers. (1st, Ed.) New Zealand, USA: Chandos

[8] Evans, N., & Price, J. (2012). Barriers to the Effective Deployment of Information Assets: An Executive Management Perspective. Interdisciplinary Journal of Information, Knowledge, and Management, 7, 23. Retrieved from www.webcitation.org/6x3CxuD6j.

[9] Fleckenstein, M., & Fellows, L. (2018). Modern Data Strategy (1 ed.). Springer International Publishing. Retrieved from Springer: https://www.springer.com/

[10] Weber, K., Otto, B., & Osterle, H. (2009). One Size Does Not Fit All - A Contingency Approach to Data Governance. Journal of Data and Information Quality, 27. doi:10.1145/1515693.1515696

[11] The Data Governance Institute. (2015). Definitions of Data Governance. Retrieved from The Data Governance Institute: www.datagovernance.com/adg_data_governance_definition/

[12] Ladley, J. (2012). Data Governance (1st ed.). Morgan Kaufmann.

[13] Ladley, J. (2019). Data Governance (2nd ed.). Academic Press.

[14] Carnegie Mellon University Software Engineering Institute SEI. (2016). History of Innovation at the SEI. USA: Software Engineering Institute-Carnegie Mellon University. Retrieved from: https://resources.sei.cmu.edu/asset_files/Book/2016_014_001_634914.pdf.

[15] Manjrekar, M. (2012, 07 16). Conducting an Effective Process Maturity Assessment: du Telecom Case Study. Retrieved from Process Excellence Network: www.processexcellencenetwork.com/innovation/articles/process-maturity-assessment

[16] User. (2018, December 4). Capability Maturity Model (CMM)-CIO Wiki. Retrieved November 18, 2019, from CIO Wiki: https://cio-wiki.org/wiki/Capability_Maturity_Model_(CMM)

[17] IGI Global. (2017). Computer systems and software engineering: concepts, methodologies, tools, and applications. USA: IGI Global. Retrieved from https://books.google.tn/.

[18] CMMI Institute. (2019). CMMI Institute LLC. Retrieved from CMMI Institute: https://cmmiinstitute.com/

[19] Lee, J., Lee, D., & Kang, S. (2007). An Overview of the Business Process Maturity Model (BPMM). In K. C.-C. Chang, W. Wang, L. Chen, -C. Ellis, C.-H. Hsu, A. Tsoi, & H. Wang (Ed.), Advances in Web and Network Technologies, and Information Management. Lecture Notes in Computer Science. 4537, pp. 384-395. Huang Shan, China: Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-540-72909-9_42

[20] Data Crossroads. (n.d.). DM Maturity. Retrieved from datacrossroads.nl.

[21] Khoshgoftar, M., & Osman, O. (2009). Comparison of maturity models. 2009 2nd IEEE International Conference on Computer Science and Information Technology, (pp. 297-301). Beijing. DOI:10.1109/ICCSIT.2009.5234402

[22] Diogo Proença, Ricardo Vieira, José Borbinha, Pável Calado, Bruno Martins. (2017, January 23). A Maturity Model for Information Governance. Lisbon, Portugal.Charles M. Dollar, Lori J. Ashley. (2012). Digital Preservation Capability Maturity Model, Version 2.4. Accessed November 2012

[23] CMMI® Institute. (2019). DATA MANAGEMENT MATURITY (DMM) MODEL. CMMI® Institute.

[24] EDM Council. (n.d.). DCAM Assessments & Support. Retrieved from EDMCouncil.org: https://edmcouncil.org/page/dcamassessmentsuppor

[25] EDM Council. (2019). DCAM v2.1.1 Foreword & Introduction. EDM Council, Inc.

[26] Christine Anderson, William Ellerby. (2018, February). Digital Maturity Model. Retrieved from Deloitte LLP: www2.deloitte.com

[27] Deloitte LLP. (2018, August). Deloitte Digital Maturity Model. Retrieved from Tmforum: www.tmforum.org

[28] Gartner, I. (2014). blogs.gartner.com. Retrieved from gartner.com: blogs.gartner.com/andrew_white/files/2016/10/On_site_poster.pdf

[29] Jennings, M. (2007, March). Developing a Roadmap for an Enterprise Information Management Program, 1. (EIMI) Retrieved from EIMInstitute.org: www.eiminstitute.org/library/eimi-archives/volume-1-issue-1-march-2007-edition/enterprise-information-management-primer

[30] Schmarzo, B. (2014, January 22). “Big Ideas” Big Data Business Model Maturity Index. Retrieved from InFocus: infocus.delltechnologies.com

[31] DAM. (2017). DAM Maturity Model. Retrieved from dammaturitymodel.org: https://www.dammaturitymodel.org/

[32] Crowston, Kevin, Jian, Qin. (2011). A Capability Maturity Model for Scientific Data Management: Evidence from the Literature. (J. Qin, Ed.) ResearchGate, 48, 9.

[33] De Simoni, G. (2020, February 24). Implement Enterprise Metadata Management to Drive Effective Enterprise Information Management. (I.a. Gartner, Ed.) Retrieved from gartner.com: https://www.gartner.com/en/documents/3981254

[34] IBM. (n.d.). Industry Data Models. Retrieved from IBM.com: www.ibm.com/analytics

[35] Firican, G. (2018, August 8). Data governance maturity models – IBM. Retrieved from LightsOnData.com: https://www.lightsondata.com/data-governance-maturity-models-ibm/

[36] Devlin, D. (2019). Models of Digital Transformation-The role of context, governance, integration, and industry models. IBM CORPORATION. 9sight Consulting.

[37] Firican, G. (2018, August 29). Data governance maturity models – Stanford. Retrieved from LightsOnData.com: www.lightsondata.com/data-governance-maturity-models-stanford/

[38] TDWI. (2014). TDWI Transforming Data with Intelligence. Retrieved from TDWI Research & Resources: tdwi.org

[39] Firican, G. (2018, September 19). Data governance maturity models - DataFlux. Retrieved from www.lightsondata.com.

[40] IBM. (n.d.). IBM Government Digital Transformation Maturity Model. Retrieved from IBM.com: www.ibm.com/industries/government/resources/digital-transformation-evaluation

[41] Devlin, D. (2019). Models of Digital Transformation-The role of context, governance, integration, and industry models. IBM CORPORATION. 9sight Consulting.

[42] István Alföldi, Zsuzsanna Fülöp, Zoltán Szatucsek. (2014, July 31). E-ARK General Model. European Archival Records and Knowledge Preservation, 1.3.

[43] Paulk, Mark C; Curtis, Bill; Chrissis, Mary Beth; Weber, Charles V. (1993). Capability Maturity Model (SM) for Software, Version 1.1. Software Engineering Institute. Pittsburgh: CMU/SEI-93-.

[44] User. (2018, December 13). Enterprise Information Management (EIM). Retrieved December 26, 2018, from cio-wiki.org.

[45] White, S. K. (2018, March 16). What is CMMI? A model for optimizing development processes. Retrieved from https://www.cio.com/article/2437864/process-improvement-capability-maturity-model-integration-cmmi-definition-and-solutions.html.

[46] Proença, D., & Borbinha, J. (2018, September). Maturity Models for Data and Information Management. doi:10.1007/978-3-030-00066-0_7

[47] Kolukısa - Tarhan, Ayça & Turetken, Oktay & Ilisulu, Fadime. (2015). Business Process Maturity Assessment: State of the Art and Key Characteristics. 10.1109/SEAA.2015.50

[48] Charles M. Dollar, Lori J. Ashley. (2012). Digital Preservation Capability Maturity Model, Version 2.4. Accessed November 2012

[49] Preservica Software company. (2014, April). Digital Preservation Maturity Model. Retrieved from: www.preservica.com.

[50] Katuu, S. (2013). The Utility of Maturity Models - The ECM Maturity Model within a South African context. Department of Information Science, University of South Africa, South Africa.

[51] Gable, J. (2015). The Principles, IG Maturity Model: Tools for Professional Growth. ARMA Magazine, 3.

[52] ARMA International. (2013). ARMA International’s Information Governance Maturity Model.

[53] ARMA International. (2017). The Principles® Maturity Model. Retrieved from ARMA International.

APPENDIX I. MATURITY MODELS OVERVIEW

Name of Family	Models' details				
	Model	***Abbreviation***	***Date***	***Author***	***Ref.***
Data Management	DAMA-DMBOK Data Management Maturity Model	DAMA-DMBOK2	2009	DAMA International	[7]
	DataFlux Master Data Management Model	MDM	2010	DataFlux Company	[44]
	Data Management Maturity Model	DMM	2014	CMMI Institute	[26]
	Data Management Capability Assessment Model	DCAM v2	2019	The Enterprise Data Management Council	[43] [33]
	The "Orange" Data Management Maturity Model	DMMM	2019	Data Crossroad	[20]
Data/Information Governance	IBM Data Governance Council Maturity Model	DGMM	2007	IBM	[33] [34] [35]
	DataFlux Data Governance Maturity Model	DGMM	2007	DataFlux Company	[44]
	The Principles Maturity Model	Principles MM	2007	ARMA International	[52]
	Stanford Data Governance Maturity Model	DGMM	2011	Stanford University's Data Governance Office	[37]
	Gartner's Enterprise Information Management Maturity Model	EIM	2016	Gartner	[29] [30] [31]
	E-ARK Information Governance Maturity Model	A2MIGO	2017	E-ARK	[42]
Software Development	Capability Maturity Model	CMM	1986	Software Engineering Institute of Carnegie- Mellon	[16]
	Capability Maturity Model Integration	CMMI	2018	Capability Maturity Model Institute	[18]
Digital Assessment	Digital Preservation Maturity Model	DPMM	2014	Preservica	[48]
	Digital Preservation Capability Maturity Model	DPCMM	2015	Preservica	[25]
	Digital Assets Management Maturity Model	DAM	2017	DAM Foundation	[44]
	Deloitte Digital Maturity Model	DMM	2018	Deloitte	[26][33]
Analytics	● TDWI: Analytics Maturity Model & Assessment ● TDWI: Self-service Analytics Maturity Model ● TDWI: IoT Data Readiness Assessment ● TDWI: Advanced Analytics Maturity Model ● TDWI: Hadoop Readiness Assessment	TDWI	2018	TDWI - Transforming Data With Intelligence	[38]
Business Performance	ECM Maturity Model	ECM3	2009	ECM	[49]
	Big Data Business Maturity Model Index	BDBM	2014	DELL	[32]