

# Establishing a Data Analytics Stage Model (DASM): Integrating “Open Coding” and Conceptual Mapping for Enhanced Maturity Assessment

Shuo Yan <sup>1,\*</sup>  
Jeff Jones <sup>2</sup>

\*Corresponding author

<sup>1</sup> PhD candidate, Warwick Manufacturing Group, University of Warwick, Coventry, UK, christina-yan.yan@warwick.ac.uk

<sup>2</sup> Warwick Manufacturing Group, University of Warwick, Coventry, UK, j.a.jones@warwick.ac.uk

## ABSTRACT

This research presents a novel Data Analytics Stage Model (DASM) to address the shortcomings of existing models. By analysing 26 prior models, several key gaps were identified: limited model synthesis, lack of methodological rigor, absence of legal compliance considerations, and unclear descriptions. DASM offers a comprehensive framework to help organisations assess their data analytics (DA) maturity, utilising transition tables to understand their developmental path. DASM's standardised structure and clear definitions provide a valuable tool for organisations seeking to enhance their DA capabilities.

**Keywords:** Data management, data analytics, CMMS, capability maturity model, stage model, digital transformation.

## INTRODUCTION

### Introduction and Trend of Digital Transformation

Digital transformation denotes the transition of enterprises, organisations, or societies from traditional non-digital or semi-digital processes to fully digitalised operations (Baranauskas & Raišienė, 2022; Boyle et al., 2022). This transition often entails leveraging digital technologies to redefine business processes, organisational culture, and customer experiences to meet the evolving demands of the market, business landscape, and environmental changes (Boyle et al., 2022; Forum, 2018; Hai et al., 2021; Merkas, 2022; Saputra et al., 2018). A 2023 KPMG study found that 63% of organisations saw positive results from digital initiatives, with many reporting profit increases over 10% (Soni, 2023). Numerous business reports estimate the market size and growth of digital transformation with a compound annual growth rate (CAGR) between 20% and 30%, indicating a positive outlook for the adoption of digital technologies in business (*Digital Transformation Market Size & Share Analysis - Industry Research Report - Growth Trends - 2032*, 2023; *Digital Transformation Market Size, Share, Growth, Trends [2032]*, 2024; *Digital Transformation Market Size, Trends Report, 2023-2030*, 2023). While digital transformation is becoming a global focus, BCG reports that only 20% of public sector initiatives have achieved their goals, compared to a 30% success rate in the private sector (*Digital Government Transformation in the Middle East*, 2022). KPMG also emphasises the need for organisations to streamline their digital transformation efforts to maximize returns (Soni, 2023). Effective data management (DM) is crucial for successful digital transformation. Organisations must prioritise data management to ensure their digital initiatives are effective.

### Introduction and Significance of Data Management and Data Analytics

DM is defined as a process that encompasses data acquisition, data quality control, metadata management, data storage and preservation, data sharing, data analytics (DA), and digital visualisation (Abraham et al., 2019; Baolong et al., 2018; Hernandez-Hall, 2021). This process covers the entire data life cycle and includes the budget allocation and corresponding infrastructure purchase required to complete this cycle (Baolong et al., 2018; Christopherson et al., 2020; Qin et al., 2014; Thabet et al., 2022). Effective DM ensures the accuracy, completeness, consistency, compliance, and security of data, allowing for high-quality data to be used to support organisations in making informed business strategies and setting sustainable development goals (Abebaw Fekadu et al., 2021; Al-Ruithe & Benkhelifa, 2017; Merkas, 2022). In this paper the specific focus is on DA as one crucial component of DM (Aras Okuyucu & Yavuz, 2020; Grossman, 2018; Hornick, 2020).

DA involves the use of statistics, algorithms, and machine learning techniques to inspect, clean, transform, and model data, aiming to uncover valuable insights, draw conclusions, and support decision-making (Aras Okuyucu & Yavuz, 2020; Grossman, 2018; Hornick, 2020). According to the definition of the data life cycle, data acquisition, metadata management, storage and preservation, and data sharing are foundational to DA (Baolong et al., 2018; Christopherson et al., 2020; Qin et al., 2014; Thabet et al., 2022). Hence, an organisation's DA capability can also indirectly reflect the completeness of its DM. Therefore, this paper utilises the Capability Maturity Model (CMM) concept to divide DA capabilities into six distinct stages.

### Introduction and Application of CMMs in Data Management

CMM, was first introduced in Watts Humphrey's "Managing the Software Process" in 1989 (Humphrey, 1989) and the concept of maturity serves as a crucial metric in evaluating an organisation's development and preparedness, spanning a spectrum from immaturity to advanced maturity (Humphrey, 1989). Maturity models offer a structured approach to pinpoint the strengths and potential areas of growth within a specific dimension or organisation (Humphrey, 1989). By leveraging these models,

organisations can derive essential insights into their existing competencies and pinpoint areas needing enhancement (Hernandez-Hall, 2021). These structured frameworks encapsulate industry best practices, enabling organisations to measure their efficacy against recognised benchmarks (Klievink et al., 2017; Vesset et al., 2015).

In recent years, many researchers have studied CMMs for DA. This interest stems from the growing number of companies seeking to improve their data management capabilities but unsure about the best approach (Qin et al., 2014; Vesset et al., 2015). To address this, researchers developed CMMs to help organisations assess their DA maturity and plan for future development. However, there are shortcomings existed in the existing CMMs, and this research has concluded the shortcomings in the next section and improved the shortcomings by establishing a new and a comprehensive model. Consequently, the main aim of this paper are:

- *To establish a Data Analytics Stage Model (DASM), which provides a standardised, structured, and objective roadmap for organisations to determine their DA maturity stage in DM to guide organisations in progressing to the next maturity stage by using this stage model as a framework.*

To achieve this main aim, the objectives will be introduced at the end of the next section.

### LITERATURE REVIEWS AND RESEARCH GAPS

Through a ten-year Systematic Literature Review (SLR) and tracking the reference lists of all pertinent papers, the author selected 26 CMMs that are fully relevant to DM or DA. In the following Table 1, the author lists these papers in the order of publication dates.

Table 1 The 26 previous DA maturity models in the last decade

Paper Name	Citation
Business Intelligence Maturity: The Economic Transitional Context Within Slovenia	(Lukman et al., 2011)
Towards A Business Analytics Capability Maturity Model	(Cosic, Shanks and Maynard, 2012)
Using Quantitative Analyses to Construct A Capability Maturity Model for Business Intelligence	(Raber, Winter and Wortmann, 2012)
A Capability Maturity Model for Research Data Management	(Qin, Crowston and Kirkland, 2014)
Big Data Maturity: An Action Plan for Policymakers and Executives	(El-Darwiche et al., 2014)
TDWI Analytics Maturity Model Guide	(Halper and Stodder, 2014)
IDC Maturity Scape: Big Data and Analytics 2.0	(Vesset et al., 2015)
Big Data Maturity Model for Malaysian Zakat Institutions to Embark on Big Data Initiatives	(Sulaiman, Cob and Ali, 2015)
Hortonworks Big Data Maturity Model	(Dhanuka, V., 2016)
How organisations leverage Big Data: a maturity model	(Comuzzi and Patel, 2016)
How Can SMEs Benefit from Big Data? Challenges and a Path Forward	(Coleman et al., 2016)
A Value Based Big Data Maturity Model	(Farah, 2017)
Big Data in The Public Sector: Uncertainties and Readiness	(Klievink et al., 2017)
Defining Analytics Maturity Indicators: A Survey Approach	(Lismont et al., 2017)
A Conceptual Framework for Assessing an Organisation's Readiness to Adopt Big Data	(Olszak and Mach-Król, 2018)
A Framework for Evaluating The Analytic Maturity of An Organization	(Grossman, 2018)
A Health Data Analytics Maturity Model for Hospitals Information Systems	(Carvalho et al., 2019)
A Data Governance Maturity Evaluation Model for government departments of the Eastern Cape province, South Africa	(Herselman, Wayi and Olaitan, 2019)
Maturing Research Data Services and The Transformation of Academic Libraries	(Cox et al., 2019)
Towards A Maturity Model for Big Data Analytics in Airline Network Planning	(Hausladen and Schosser, 2020)
A Data Science Maturity Model for Enterprise Assessment	(Hornick, 2020)
Nonprofit Data Management: A Stage Model	(Hernandez-Hall, 2021)
Data Management Maturity Model (DMMM)	(Zitoun et al., 2021)
Evaluation of Enterprise Accounting Data Management Based on Maturity Model	(Xiang, 2022)
The Development Of The Data Science Capability Maturity Model: A Survey-Based Research	(Gökulp et al., 2022)

Paper Name	Citation
Data Science Maturity Models (DSMMs): A Broad Consensus Composite	("Data Science Maturity Models (DSMMs)," 2023)

The twenty-six listed papers have been critically reviewed and compared based on (1) their methodological approach to model development, (2) the comprehensiveness of their coverage from the perspective of legal compliance, and (3) the clarity and comprehensiveness of their descriptions of dimensions, subdimensions, stages, and transitions.

In their approach to model development, only four maturity models incorporated or synthesised existing models for their development after conducting a comparative analysis of available models (Cosic et al., 2012; Hausladen & Schosser, 2020; Hernandez-Hall, 2021; Lukman et al., 2011). Among these, only one study integrated more than 10 models and applied a specific method—Thematic Content Analysis—for this synthesis (Cosic et al., 2012). Although Hausladen and Schosser (2020) referenced 19 maturity models, they utilised only six in their model creation. Notably, none of these studies employed Grounded Theory for the integration of models, despite its suitability for developing new frameworks or theories through systematic and iterative methods (Charmaz, 2006; Noble & Mitchell, 2016; Soliman & Kan, 2004). Additionally, none of the reviewed models used Conceptual Mapping as part of their synthesis approach. From the standpoint of legal compliance, none of the existing models addressed aspects related to data security, privacy, or ethics comprehensively. Regarding the clarity and comprehensiveness of dimensions, subdimensions, stages, and transitions, nine of the 26 models included transition tables. However, only five models provided clearly defined, comprehensive, and consistent descriptions (Comuzzi & Patel, 2016; *Data Science Maturity Models (DSMMs)*, 2023; Hausladen & Schosser, 2020; Qin et al., 2014; Vesset et al., 2015).

Therefore, the objectives of this paper are to:

- To establish a comprehensive DASM by synthesising 26 models from the last decade;
- To emphasise the use of Conceptual Mapping to establish a standardised structure for the DASM;
- To demonstrate transparent processes of utilising Grounded Theory of "open coding";
- To ensure the DASM covers compliance with data laws related to data privacy, data security, and data ethics;
- To create a DASM with detailed descriptions of transitions tables between stages.

## METHODOLOGY

To develop a comprehensive DASM, this research follows two key steps: 1) applying Conceptual Mapping to establish an initial structure of the DMSM based on the analysis of 14 selected papers, and 2) synthesising all 26 DA maturity models using Grounded Theory.

Conceptual Mapping, implemented using MindManager (*Support*, 2024), was employed to visualise and create the initial DASM structure. This technique helps organise and represent the relationships and themes found in the literature by creating visual diagrams or maps, which summarise complex research landscapes (Moon et al., 2011; Rosas & Ridings, 2017; Verdonck et al., 2015). It supports the iterative refinement of maturity models by mapping connections between concepts and theories (Schumacher, 2015; Schumacher et al., 2016; Steinlechner et al., 2021; Zoubek & Šimon, 2021). As shown in Figure 1, which is an example of the mapping processes of the Stage 3 of DASM, and Figure 2 shows the color of each paper referred in Figure 1. The yellow colour in the map corresponds to concepts derived from paper 2, El-Darwiche *et al.* (2014) which stated that “Involve sales and marketing techniques such as customer segmentation and targeting, or entry-stage analytic methods for product recommendation”, is identified as a description of Subdimension of Ability to Perform of Dimension of Data Analytics in Stage 3. Each line within the maturity models published in the 14 papers was placed in the appropriate position in the contextual map, thus defining the structure and descriptions of the stages, dimensions, and subdimensions at this initial stage. This example was chosen because it is the most complex stage with the largest number of references.

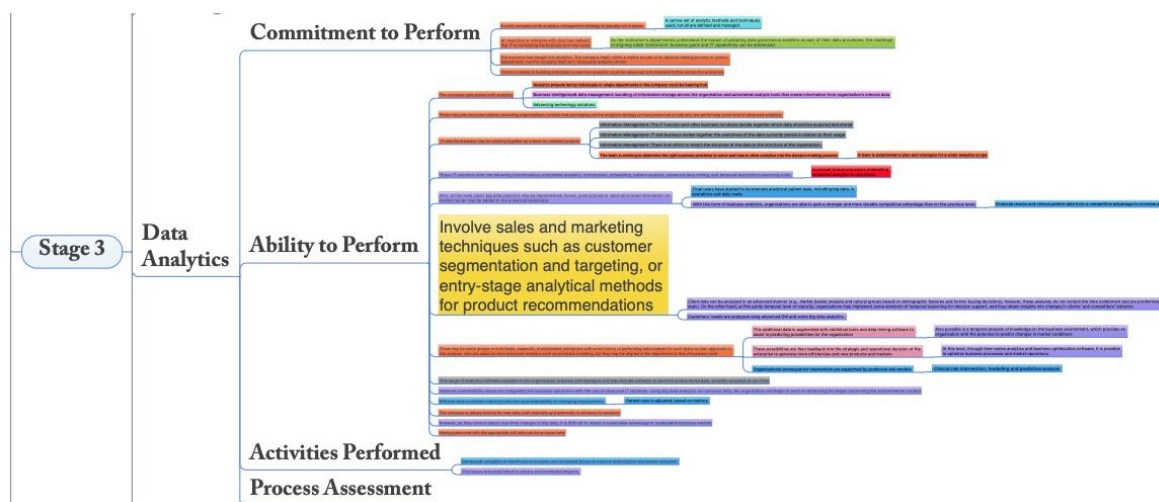


Figure 1: Conceptual Mapping processes of Stage 3 of DASM

Paper NO.	Colour REF.	Paper Name	Year	Author
1		A Capability Maturity Model for Research Data Management	2014	Qin, Crowston and Kirkland
2		Big Data Maturity: An Action Plan for Policymakers and Executives	2014	El-Darwiche et al.
3		TDWI Analytics Maturity Model Guide	2014	Halper and Stodder
4		IDC MaturityScape: Big Data and Analytics 2.0	2015	Vesset et al.
5		Big Data Maturity Model for Malaysian Zakat Institutions to Embark on Big Data Initiatives	2015	Hidayah Sulaiman, Cob and Ali
6	None	Hortonworks Big Data Maturity Model	2016	Vishal Dhanuka
7		How organisations leverage Big Data: a maturity model	2016	Comuzzi and Patel
8		A Value Based Big Data Maturity Model	2017	Farah
9		Big data in the public sector: Uncertainties and readiness	2017	Klievink et al.
10		A Conceptual Framework for Assessing an Organization's Readiness to Adopt Big Data	2018	Olszak and Mach-Król
11		A health data analytics maturity model for hospitals information systems	2019	Carvalho et al.
12		A Data Governance Maturity Evaluation Model for government departments of the Eastern Cape province, South Africa	2019	Olaitan, Herselman and Wayi
13		NONPROFIT DATA MANAGEMENT: A STAGE MODEL	2021	Ashley M Hernandez-Hall
14		Data Management Maturity Model (DMMM)	2021	Zitoun et al.

Figure 2: Reference Colour Bar for the Contextual Mapping

In line with Conceptual Mapping, Grounded Theory of “open coding” also called “line-by-line coding” was used in this research to complete and refine the initial DASM structure (Barney & Anselm, 2017). Grounded Theory, introduced by Glaser and Strauss in “Awareness of Dying” (1965), uses the constant comparative method to iteratively develop theories (Andrews & Nathaniel, 2015; Barney & Anselm, 2017; Glaser & Strauss, 2017). It has been applied in the creation of maturity models to identify key factors, dimensions, and stages (Geyer & Krumay, 2015; Klötzer & Pflaum, 2017; Mahdavi et al., 2022). In the context of establishing maturity models, it facilitates the identification and comprehension of the underlying factors, dimensions, and stages of maturity within specific domains. Grounded Theory studies typically involve iterative processes of data collection and analysis, enabling researchers to continuously refine their theories based on emerging insights (Chun Tie et al., 2019). In this study, Grounded Theory was applied to establish the DASM using Nvivo software (*NVivo – Top Preise, Training Und Support*, 2024). By synthesising each line of the 26 existing maturity models iteratively, as shown in Figure 3 in the next section, the final structure of the DASM with defined dimensions were established, and the descriptions of each diemnsion in each stage was assigned. In this amalgamation process a number of assumptions were made:

- *Assumption 1: If a concept is only mentioned at a higher maturity stage, there is an absence of this concept in the earlier stages*
- *Assumption 2: In statements about the maturity of DM where DA is not mentioned explicitly, DA has the same maturity level as DM, if the papers' definition of DM includes DA.*

Same to the example mentioned above, the synthesis processes of Stage 3 of DASM is outlined below as an example.

### Example of Stage 3 Synthesising Processes by Using Grounded Theory

#### Strategy Planning

A strategy for exploring DA trends and tools has been developed in specific departments, such as the IT department (*Data Science Maturity Models (DSMMs)*, 2023; Gökalp et al., 2022; Halper & Stodder, 2014; Hausladen & Schosser, 2020, 2020; Hornick, 2020; Lismont et al., 2017; Lukman et al., 2011; Raber et al., 2012; Zitoun et al., 2021). Concurrently, a goal of increasing profits through Big Data Analytics (BDA) has been set (Zitoun et al., 2021).

#### Data Governance

The organisation has considered some DA policies in alignment with its current business objectives (Cotic et al., 2012; Cox et al., 2019; *Data Science Maturity Models (DSMMs)*, n.d.; Hausladen & Schosser, 2020; Hornick, 2020, 2020; Raber et al., 2012; Sulaiman et al., 2015). However, there is no steering committee in place to oversee the DA process, and metrics to assess the quality of DA have not been established (Comuzzi & Patel, 2016; *Data Science Maturity Models (DSMMs)*, 2023; Vesset et al., 2015; Xiang, 2022).

### **Leadership and Culture**

The organisation's leadership recognises the importance of DA and encourages its staff to explore new analytical methods and techniques (Carvalho et al., 2019; *Data Science Maturity Models (DSMMs)*, 2023; Gökalp et al., 2022; Halper & Stodder, 2014; Hausladen & Schosser, 2020; Hornick, 2020; Sulaiman et al., 2015; Zitoun et al., 2021). The leadership has supported and encouraged the utilisation of data analytic methods and techniques, and at least one executive sponsor of BDA application at this stage (Halper & Stodder, 2014; Hausladen & Schosser, 2020). While these methods are being adopted for specific tasks, projects, or departments, they are not yet integrated into the broader decision-making processes (*Data Science Maturity Models (DSMMs)*, 2023; Halper & Stodder, 2014; Hausladen & Schosser, 2020). Although most employees and stakeholders see the value in DA, there is still some resistance or lack of understanding (Halper & Stodder, 2014; Hausladen & Schosser, 2020; Hornick, 2020). A noticeable gap exists in the collaboration between the IT department and other business departments (*Data Science Maturity Models (DSMMs)*, 2023; Halper & Stodder, 2014; Hornick, 2020). Even though IT remains the central hub for DA, other departments have continued actively posing strategic questions that they expect the IT team to address (*Data Science Maturity Models (DSMMs)*, 2023; Halper & Stodder, 2014). Also, there might be some ad-hoc collaborations between the IT departments and other business departments, and this especially includes interactions with data analysts and scientists from various departments (Hornick, 2020). Such collaborations typically involve sharing and modifying data science assets like data sets, visualisation plots, graphs, codes in R, Python, SQL, and scripts in web-based notebooks like Zeppelin and Jupyter. However, the tools and platforms used for these collaborations are not standardised and consistent across the entire organisation (Hornick, 2020).

### **Ability to Perform**

The organisation has improved its capability to deeply drive sights from data, and some of the DA projects led by individuals or departments are starting to show results and could be used for operational decision-making (Halper & Stodder, 2014; Olszak & Mach-Król, 2018). The role of data scientist has been introduced to harness the potential of machine learning and other advanced techniques (Hornick, 2020). In addition to the pre-existing DA competencies and tools, there has been an incorporation of BDA techniques and tools (Hausladen & Schosser, 2020; Herselman et al., 2019; Hornick, 2020). Therefore, most tactical operational optimisation decisions have been automated (Hausladen & Schosser, 2020).

Most structured data and limited unstructured data have been analysed (Farah, 2017; Halper & Stodder, 2014; Hausladen & Schosser, 2020; Zitoun et al., 2021). Both time-stamped data and sequential data are collected and used for DA (Olszak & Mach-Król, 2018); however, the organisation cannot detect the real-time changes at this stage (Hausladen & Schosser, 2020; Olszak & Mach-Król, 2018). DA processes leverage a mix of both internal and external data sources (Carvalho et al., 2019; Comuzzi & Patel, 2016; Grossman, 2018; Hausladen & Schosser, 2020; Hornick, 2020). Descriptive DA (e.g., to generate more efficient and consistent daily productivity reports), diagnostic analytics (e.g., to do time-stamped descriptive reasoning) and limited predictive analytics (e.g., to do time-stamped pattern analytics) are conducted (Carvalho et al., 2019; *Data Science Maturity Models (DSMMs)*, 2023; El-Darwiche et al., 2014; Grossman, 2018; Hornick, 2020; Olszak & Mach-Król, 2018). For instance, to build, validate, and update models such as regression models and decision making trees (Grossman, 2018). Analytical activities are consolidated for specific projects, with collaboration seen between the IT department and other business sectors for these endeavours (Gökalp et al., 2022; Halper & Stodder, 2014; Hausladen & Schosser, 2020; Hornick, 2020). Professionals experienced in DA may be grouped into dedicated teams, focusing on addressing pivotal business challenges and offering recommendations for the broader application of DA tools (*Data Science Maturity Models (DSMMs)*, 2023; Halper & Stodder, 2014; Hornick, 2020). Additionally, the organisation might engage external consultancies equipped with more advanced tools to tackle specific business challenges (Halper & Stodder, 2014; Hausladen & Schosser, 2020). Investments have been channelled towards state-of-the-art analytics, post the experimental phase, with prototypes being utilised for testing and refining (*Data Science Maturity Models (DSMMs)*, 2023; Farah, 2017; Gökalp et al., 2022; Halper & Stodder, 2014; Hausladen & Schosser, 2020; Hornick, 2020).

It is anticipated that CDOs will take the lead in framing the data strategy, emphasising data-driven decisions with a clear business orientation and supervising data analysts (Hornick, 2020). While DA skills and tasks aren't explicitly mentioned in job descriptions, there is either a continuous effort or plans in the pipeline to revisit or adjust this, primarily due to the organisation's challenge of a skills gap (Hausladen & Schosser, 2020; Hernandez-Hall, 2021). There's also an emphasis on continual learning, with on-the-job training provided for both specific and general tasks (Farah, 2017; Hernandez-Hall, 2021; Hornick, 2020). The organisation promotes embedded analytics, eliminating the need for staff to rely on separate applications for their DA needs (Olszak & Mach-Król, 2018; Zitoun et al., 2021). The performance of employed infrastructure is satisfactory, which almost meets expectations but is not exceptional (Cosic et al., 2012; Cox et al., 2019; Gökalp et al., 2022; Hausladen & Schosser, 2020; Lukman et al., 2011; Vesset et al., 2015).

### **Activities Performed**

Some specialised procedures and processes related to DA have been codified for ad-hoc projects, and some standardised procedures and processes have been defined throughout the entire organisation (Carvalho et al., 2019; Comuzzi & Patel, 2016; Gökalp et al., 2022; Hausladen & Schosser, 2020; Hernandez-Hall, 2021; Hornick, 2020; Raber et al., 2012; Sulaiman et al., 2015).

### Process Assessment

The approaches to MAV the procedures and processes related to DA have not been defined but might be considered at this stage (Gökalp et al., 2022; Grossman, 2018; Hausladen & Schosser, 2020; Hornick, 2020; Vesset et al., 2015; Xiang, 2022).

### DATA ANALYTICS STAGE MODEL (DASM)

By using the approaches introduced in the last section, the final structure and the transitions between stages of DASM have been established, which are introduced below.

#### A Holistic Structure and Definitions Related to DASM

The DASM comprises six stages numbered from 0 to 5. Higher stages indicate a more mature DA capability. Specifically, Stage 0 pertains to organisations with minimal or zero DA capabilities, which is often characteristic of very small start-up businesses. In contrast, Stage 5 represents the pinnacle of maturity with advanced DA capability.

To assess an organisation's current DA Stage, six dimensions are used: Strategy Planning, Data Governance, Leadership and Culture, Ability to Perform, Activities Performed, and Process Assessment. By evaluating these dimensions, organisations can discern not only their current position but also their potential future trajectories. As depicted in Figure 3 below, the definitions for each dimension are provided in a checklist format.

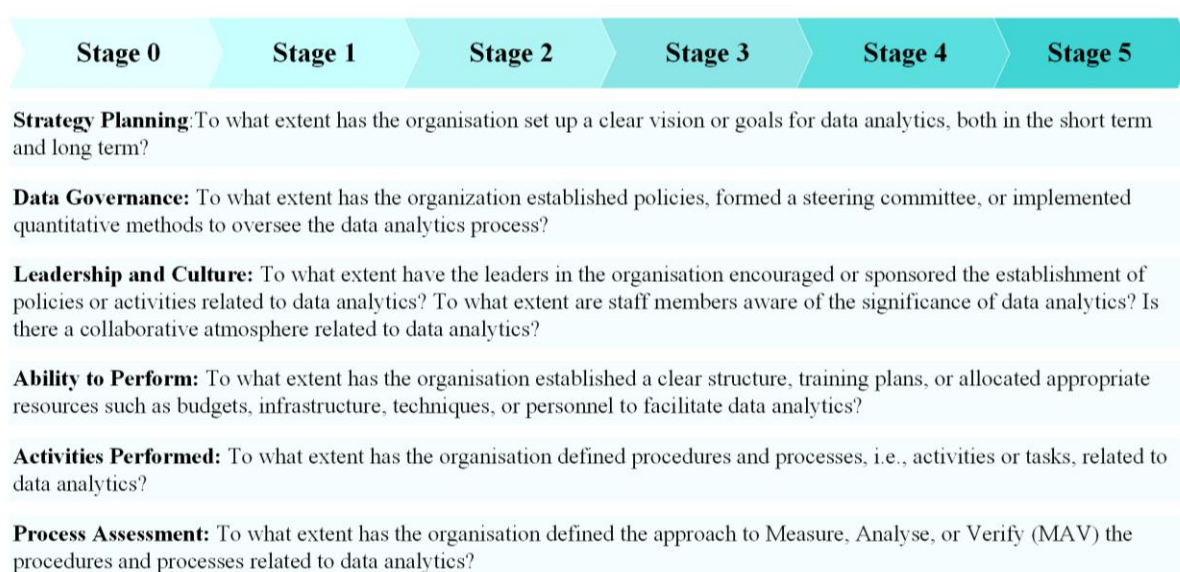


Figure 2: The structure of DASM

#### The Transition Tables of DASM Across Dimensions and Stages

The transition tables of DASM across different dimensions and stages are outlined below in Table 2. By using the table, the users can easily locate their current maturity of the certain capability below the certain dimension, and the way to move forward could also be gained.

Table 2 DASM transition tables

Dimension	Capability	Stage 0	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
Strategic Planning	Strategies related to the exploration of data analytics trends and tools:	Have not been considered.	Have been considered by certain individuals.	Have not been defined, however, the organisation has started assessing its current capability of data analytics for seeking projects.	Have been defined in certain department(s).	Have been defined in the entire organisation.	Have been defined as aligning with self-service deployment in the entire organisation.
	Goals establishment related to data analytics:	There are no goals established.			There is a goal of making more profits by utilising big data analytics.	There is a goal of making more profits by utilising big data analytics and being sustainable by leveraging big data analytics.	There is a goal of making more profits, and being sustainable, customer-oriented and forward-thinking by



							leveraging big data analytics.
Dimension	Capability	Stage 0	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
<b>Data Governance</b>	<b>Policy establishment related to data analytics:</b>	Have not been considered.			Some policies have been addressed to align with the current business objectives of the organisation.	All policies have been established throughout the entire organisation, including standardised and clear guidelines for most of the analytical methods and techniques in the organisation.	All policies are continuously improved through regular evaluation, including the guidelines of self-service deployment.
	<b>The existence of a steering committee related to data analytics:</b>	There is no steering committee.				There might be some certain individuals or department(s) to assess the data analytics quality.	There is a formalised steering committee for data analytics.
Dimension	Capability	Stage 0	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
<b>Leadership and Culture</b>	<b>Leadership in data analytics methods, techniques, and tools:</b>	There is a lack of leadership.	Leaders have recognised the importance and benefits of data analytics.	Leaders have encouraged staff to explore new analytical methods and techniques, and they have merged to carry out the discussion about it.	The leadership have sponsored staff to explore new analytical methods and techniques.	The leadership have sponsored and underscored the significance of data analytics as a fundamental value and competitive advantage.	The leadership have sponsored, and regularly delve into emerging analytics methods and techniques.
	<b>Attitudes to data analytics methods, techniques, and tools:</b>	Have not been considered and trusted by the staff member.	Have only been considered by certain individuals who handle data from the IT department.	Have actively been considered, trusted, valued by certain individuals.	Have actively been considered, trusted, valued by most staff and stakeholders.	Have actively been considered, trusted, valued and understood by all staff and stakeholders.	
Dimension	Capability	Stage 0	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
<b>Ability to Perform</b>	<b>1.Automation of decision making processes:</b>	There is no automation of decision-making	Basic tasks are automated, such as calculations performed using spreadsheet functions	A few number of tactical operational optimisation decisions being automated.	Most tactical operational optimisation decisions have been automated.		Most tactical and strategic operational optimisation decisions are automated.
	<b>2.Data types utilised for data analytics:</b>	Have not been considered.	Structured static data is used.	Structured static and sequential data are used.	Structured sequential and time-stamped data, and unstructured static data are used.	Structured temporal big data, and unstructured static data are used.	Structured and unstructured temporal big data are used, and the organisation mainly use unstructured data from social media.
	<b>3.Methods of data analytics:</b>	No methods, the decision-making solely based on leaders' experiences and beliefs.	Descriptive data analytics is mainly conducted by using historical data (e.g., to generate daily productivity reports).	Descriptive data analytics (e.g., to generate daily productivity reports), diagnostic analytics (e.g., to do sequential descriptive reasoning)	Descriptive data analytics (e.g., to generate more efficient and consistent daily productivity reports), diagnostic analytics (e.g., to do time-stamped descriptive reasoning) and limited predictive analytics (e.g., to do time-stamped	Descriptive data analytics (e.g., to generate temporal daily productivity reports), diagnostic analytics (e.g., to do some temporal and time-stamped descriptive reasoning) and advanced predictive analytics (e.g., to do some temporal and time-stamped	Descriptive data analytics (e.g., to generate temporal daily productivity reports), diagnostic analytics (e.g., to do temporal descriptive reasoning), predictive analytics (e.g., to do advanced data mining) and prescriptive analytics (e.g., optimisation and

					pattern analysis) are conducted.	pattern analysis) are conducted.	scheduling) are conducted.
	<b>4.Exploration of new and advanced analytics methods, techniques and tools:</b>	Have not been considered.	Have been initialised by a few individuals.	The experiments of new tools have been initialised by certain departments, and the investment for more automation tools from third parties have been considered.	The prototypes have been used to test and refine new tools, and the organisation have started investing money for it.	The prototypes have been used to test and refine new tools, and the organisation have kept investing money for it.	The organisation have kept investing more money for it and the innovative trends are continuously seek through regular investigation.
	<b>5.Roles and responsibilities related to data analytics:</b>	Have not been considered.	Have been considered by certain individuals.	Has been addressed for ad-hoc projects/certain departments.	Has been defined throughout the entire organisation.	Have been defined with quantitative quality goals.	Have been continuously improved through regular evaluation.
	<b>6.Techniques and tools utilised for data analytics:</b>	The organisation manually documents, classifies, and summarises data without the use of specialised analytical tools.	Initial data analysis primarily relies on traditional desktop spreadsheets and open-source tools.	Analytical tools and skills are limited, but R, Python, and open-source analytics techniques and tools have started being used for very important ad-hoc projects by certain experienced data analysts.	Business Intelligence (BI) systems, data mining techniques and tools, and Business Analytics (BA) systems (including the usage of R, Python, SPSS, etc.) are utilised in certain departments or are in progress for Proof of Concepts (POCs).	BI systems, most data mining techniques (such as web mining and text mining), along with in-memory processing techniques, BA systems, and the embedded analytics system have been applied throughout the entire organisation.	All possible advanced analytics techniques and advanced tools have been applied and integrated into a single application throughout the entire organisation.
<b>Dimension</b>	<b>Capability</b>	<b>Stage 0</b>	<b>Stage 1</b>	<b>Stage 2</b>	<b>Stage 3</b>	<b>Stage 4</b>	<b>Stage 5</b>
<b>Activities Performed</b>	<b>Procedures and processes related to data analytics:</b>	Have not been considered.	Have been considered by certain individuals.	Has been addressed for ad-hoc projects.	Has been defined throughout the entire organisation.	Have been defined with quantitative quality goals.	Have been continuously improved through regular evaluation.
<b>Dimension</b>	<b>Capability</b>	<b>Stage 0</b>	<b>Stage 1</b>	<b>Stage 2</b>	<b>Stage 3</b>	<b>Stage 4</b>	<b>Stage 5</b>
<b>Process Assessment</b>	<b>The approaches to MAV the processes and procedures related to data analytics:</b>	Have not been considered.	Have been considered by certain individuals.	Has been addressed for ad-hoc projects.	Has been defined throughout the entire organisation.	Have been defined with quantitative quality goals.	Have been continuously improved through regular evaluation.

## DISCUSSION AND CONCLUSION

As digital transformation advances, many organisations have made substantial investments in infrastructure and tools. However, only a few have realised the expected benefits, primarily due to ineffective data management. Over the years, various CMMs have been developed to help organisations improve their digital transition capabilities. This research reviewed and analysed 26 DA-related CMMs from the past decade, revealing that none utilised Conceptual Mapping or Grounded Theory in their model development. Moreover, none addressed legal compliance concerning data security, privacy, and ethics, and only a few provided clearly defined transition tables.

In response, this paper introduces an innovative approach to constructing a DASM, addressing the key gaps identified in the review of the 26 previous models. This approach synthesises insights from a decade's worth of DA models, with a focus on embedding a rigorous conceptual framework into the maturity modelling process. The central innovation of this model is the use



of Conceptual Mapping, which played a pivotal role in developing a standardised structure for the DASM. This method allowed for the visualisation and integration of complex concepts from different models, ensuring that the framework is both comprehensive and well-organised. Conceptual Mapping also facilitated the identification and alignment of core dimensions and stages, promoting the clarity and consistency that previous models lacked.

Additionally, Grounded Theory, specifically the "open coding" technique, has been instrumental in providing a robust empirical foundation for the DASM. This methodological approach enabled the systematic extraction and synthesis of key attributes from the reviewed models, ensuring that the DASM is built on a solid foundation of qualitative data. The transparency and rigour of Grounded Theory contributed to the detailed documentation of the model development process, enhancing the model's credibility and applicability.

The SLR process highlighted a significant research gap: the integration of data laws concerning privacy, security, and ethics. Addressing this gap within the DASM marks a substantial improvement in the model's relevance and utility in today's data-driven environments. Future research will focus on refining the transitions between compliance stages by gathering domain experts' suggestions. For example, questions such as "To what extent do you think data laws, privacy, and ethics are significant in data analytics?" and "What capabilities should be added to DASM to highlight their importance?" will be included in a survey to gather expert feedback. Also, the current DASM will be validated further by domain expert assessment by using the semi-structured survey template provided by Salah et al. (2014). The number of rounds in the expert assessment process will depend on data saturation and the validation outcomes.

Following the validation process, DASM will be tested in real-world scenarios by applying it in several organisations. These companies will conduct self-assessments of their DA maturity using the DASM's transition tables and scoring scheme to ensure its practical application.

The primary contribution of this research is the establishment of a comprehensive DASM. DASM is the most extensive DA CMM to date, integrating the dimensions and content of all related CMMs developed over the past decade. As a tool, DASM serves as a roadmap, enabling organisations to assess their current DA capabilities and determine the next steps for improvement.

## REFERENCES

- Fekadu, A., Assefa, E., Tesfaye, A., Hanlon, C., Adefris, B., Manyazewal, T., Newport, M. J., & Davey, G. (2021). Towards effective and sustainable global academic partnerships through a maturity model informed by the capability approach. *Globalization and Health*, 17, 1–7. <https://doi.org/10.1186/s12992-021-00785-2>
- Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. In *International Journal of Information Management*, 49, 424–438. <https://doi.org/10.1016/j.ijinfomgt.2019.07.008>
- Al-Ruithe, M., & Benkhelifa, E. (2017). Cloud data governance maturity model. *ACM International Conference Proceeding Series*, 151, 1–10. <http://dx.doi.org/10.1145/3018896.3036394>
- Andrews, T., & Nathaniel, A. (2015). Awareness of dying remains relevant after fifty years. *Grounded Theory Review*, 14(2), 3–10.
- Aras Okuyucu, & Yavuz, N. (2020). Big data maturity models for the public sector: A review of state and organizational level models. *Transforming Government: People, Process and Policy*, 14(4), 681–699. <https://doi.org/10.1108/TG-09-2019-0085>
- Baolong, Y., Hong, W., & Haodong, Z. (2018). Research and application of data management based on Data Management Maturity Model (DMM). *PROCEEDINGS OF 2018 10TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING AND COMPUTING (ICMLC 2018)*, 157–160. <https://doi.org/10.1145/3195106.3195177>
- Baranauskas, G., & Raišienė, A. G. (2022). Transition to Digital Entrepreneurship with a Quest of Sustainability: Development of a New Conceptual Framework. *Sustainability*, 14(3), 1104. <https://doi.org/10.3390/su14031104>
- Barney, G. G., & Anselm, L. S. (2017). *Awareness of Dying*. Routledge. <https://doi.org/10.4324/9781351327923>
- Boyle, C., Ryan, G., Bhandari, P., Law, K. M. Y., Gong, J., & Creighton, D. (2022). Digital Transformation in Water Organizations. In *Journal of Water Resources Planning and Management*, 148(7). [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0001555](https://doi.org/10.1061/(ASCE)WR.1943-5452.0001555)
- Carvalho, J. V., Rocha, Á., Vasconcelos, J., & Abreu, A. (2019). A health data analytics maturity model for hospitals information systems. *International Journal of Information Management*, 46, 278–285. <https://doi.org/10.1016/j.ijinfomgt.2018.07.001>
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. sage.
- Christopherson, L., Mandal, A., Scott, E., & Baldin, I. (2020). Toward a Data Lifecycle Model for NSF Large Facilities. In *ACM International Conference Proceeding Series*, 168–175. <https://doi.org/10.1145/3311790.3396636>
- Tie, Y.C., Birks, M., & Francis, K. (2019). Grounded theory research: A design framework for novice researchers. *SAGE Open Medicine*, 7, 2050312118822927. <https://doi.org/10.1177/2050312118822927>
- Comuzzi, M., & Patel, A. (2016). How organisations leverage Big Data: A maturity model. *Industrial Management & Data Systems*, 116(8), 1468–1492. <https://doi.org/10.1108/IMDS-12-2015-0495>
- Cosic, R., Shanks, G., & Maynard, S. (2012). *Towards a business analytics capability maturity model*. 1–11.

- Cox, A. M., Kennan, M. A., Lyon, L., Pinfield, S., & Sbaifi, L. (2019). Maturing research data services and the transformation of academic libraries. In *Journal of Documentation*, 75(6), 1432-1462. <https://doi.org/10.1108/JD-12-2018-0211>
- Galileo. (2024). Data Science Maturity Models (DSMMs): A Broad Consensus Composite. *Galileo*. Retrieved from <https://www.arcalea.com/blog/data-science-maturity-models>. (accessed 24 October 2023).
- Mourtada R, Felden, F, Bruehwiler, R. & Joykutty, T.(2022). Digital Government Transformation in the Middle East. *BCG Global*. Retrieved from <https://www.bcg.com/publications/2022/government-digital-trasformation-in-the-middle-east> (accessed 24 October 2024).
- Markets and Markets. (2024). Digital Transformation Market by Business Transformation (Process, Model, Domain), Technology (Cloud Computing, AI, Big Data Analytics(Predictive, Prescriptive), Blockchain, Cybersecurity (Network, Cloud Security), IoT, RPA) - Global Forecast to 2030. *Markets and Markets*. Retrieved from <https://www.marketsandmarkets.com/Market-Reports/digital-transformation-market-43010479.html> (accessed 24 October 2024).
- Fortune Business Insight. (2024). Digital Transformation Market Size, Share & Industry Analysis, By Technology (Cloud Computing, Big Data and Analytics, Cybersecurity, Artificial Intelligence, Internet of Things, Business Intelligence, and Blockchain), By Enterprise Type (Large Enterprises and Small & Medium Enterprises (SMEs)), By Deployment (On-Premise and Cloud), By Industry (Banking, Financial Services and Insurance (BFSI), Manufacturing, IT and Telecommunications, Retail and Consumer Goods, Healthcare, Transportation and Logistics, Government, and Others), and Regional Forecast, 2024-2032. *Fortune Business Insight*. Retrieved from <https://www.fortunebusinessinsights.com/digital-transformation-market-104878> (accessed 24 October 2024).
- Grand View Reserch. (2023). Digital Transformation Market Size, Share, Growth & Trends Analysis Report By Solution, By Deployment, By Service, By Enterprise Size, By End-use, By Region, And Segment Forecasts, 2024 - 2030. *Grand View Reserch*. Retrieved from <https://www.grandviewresearch.com/industry-analysis/digital-transformation-market> (accessed 24 October 2023).
- El-Darwiche, B., Koch, V., Meer, D., Shehadi, R., & Tohme, W. (2014). Big data maturity: An action plan for policymakers and executives. *The Global Information Technology Report*, 43, 51.
- Farah, B. (2017). A Value Based Big Data Maturity Model. *Journal of Management Policy and Practice*, 18(1), 11–18.
- Forum, W. E. (2018). *Digital Transformation Initiative Maximizing the Return on Digital Investments*.
- Geyer, S., & Krumay, B. (2015). Development of a social media maturity model—a grounded theory approach. *2015 48th Hawaii International Conference on System Sciences*, 1859–1868.
- Glaser, B., & Strauss, A. (2017). *Discovery of grounded theory: Strategies for qualitative research*. Routledge.
- Gökalp, M. O., Gökalp, E., Kayabay, K., Koçyiğit, A., & Eren, P. E. (2022). The development of the data science capability maturity model: A survey-based research. *Online Information Review*, 46(3), 547–567. <https://doi.org/10.1108/OIR-10-2020-0469>
- Grossman, R. (2018). A framework for evaluating the analytic maturity of an organization. *International Journal of Information Management*, 38(1), 45–51. <https://doi.org/10.1016/j.ijinfomgt.2017.08.005>
- Hai, T. N., Van, Q. N., & Thi Tuyet, M. N. (2021). Digital transformation: Opportunities and challenges for leaders in the emerging countries in response to COVID-19 pandemic. *Emerging Science Journal*, 5(1), 21–36. <https://doi.org/10.28991/esj-2021-SPER-03>
- Halper, F., & Stodder, D. (2014). TDWI analytics maturity model guide. *TDWI Research*, 1–20.
- Hausladen, I., & Schosser, M. (2020). Towards a maturity model for big data analytics in airline network planning. *JOURNAL OF AIR TRANSPORT MANAGEMENT*, 82, 101721. <https://doi.org/10.1016/j.jairtraman.2019.101721>
- Hernandez-Hall, A. M. (2021). *Nonprofit Data Management: A Stage Model* [PhD Thesis]. University of Nevada, Las Vegas.
- Herselman, M., Wayi, N., & Olaitan, O. (2019). A Data Governance Maturity Evaluation Model for government departments of the Eastern Cape province, South Africa. *South African Journal of Information Management*, 21(1), 1–12. <https://doi.org/10.4102/sajim.v21i1.996>
- Hornick, M. (2020). *A Data Science Maturity Model for Enterprise Assessment*. Oracle Machine Learning.
- Humphrey, W. S. (1989). *Managing the software process*. Addison-Wesley Longman Publishing Co., Inc.
- Klievink, B., Romijn, B.-J., Cunningham, S., & de Bruijn, H. (2017). Big data in the public sector: Uncertainties and readiness. *Information Systems Frontiers*, 19(2), 267–283. <https://doi.org/10.1007/s10796-016-9686-2>
- Klötzer, C., & Pflaum, A. (2017). *Toward the development of a maturity model for digitalization within the manufacturing industry's supply chain*. Hawaii International Conference on System Sciences (HICSS) 2017
- Lismont, J., Vanthienen, J., Baesens, B., & Lemahieu, W. (2017). Defining analytics maturity indicators: A survey approach. *International Journal of Information Management*, 37, 114–124. <https://doi.org/10.1016/j.ijinfomgt.2016.12.003>
- Lukman, T., Hackney, R., Popovic, A., Jaklic, J., & Irani, Z. (2011). Business Intelligence Maturity: The Economic Transitional Context Within Slovenia. *Information Systems Management*, 28(3), 211–222. <https://doi.org/10.1080/10580530.2011.585583>
- Mahdavi, M., Sheikh Al-Islami, N., Hassan Moradi, N., & Shukri, N. (2022). Design the People Capability Maturity Model among Faculty Members Based on Grounded Theory (Case Study: Islamic Azad University of Fars Province). *Journal of New Approaches in Educational Administration*, 12(6), 167–152. <https://doi.org/10.30495/jedu.2022.24472.4984>
- Merkas, Z. (2022). Digital transformation index based on the digital maturity model: Case study of croatia. *Economic and Social Development: Book of Proceedings*, 182–193.

- Moon, B., Hoffman, R. R., Novak, J., & Canas, A. (2011). *Applied concept mapping: Capturing, analyzing, and organizing knowledge*. CRC Press.
- Noble, H., & Mitchell, G. (2016). What is grounded theory? *Evidence-Based Nursing*, 19(2), 34–35. <https://doi.org/10.1136/eb-2016-102306>
- NVivo. (2024). The Most Trusted Qualitative Analysis Software (QDA) is Even Better. NVivo. Retrieved from <https://www.nvivo.de/en/> (accessed 24 October 2024).
- Olszak, C. M., & Mach-Król, M. (2018). A Conceptual Framework for Assessing an Organization's Readiness to Adopt Big Data. *Sustainability*, 10(10), 3734. <https://doi.org/10.3390/su10103734>
- Qin, J., Crowston, K., & Kirkland, A. (2014). A Capability Maturity Model for Research Data Management. *Proceedings of the American Society for Information Science and Technology*, 48(1). <https://doi.org/10.1002/meet.2011.14504801036>
- Raber, D., Winter, R., & Wortmann, F. (2012). Using Quantitative Analyses to Construct a Capability Maturity Model for Business Intelligence. *2012 45th Hawaii International Conference on System Sciences*, 4219–4228. <https://doi.org/10.1109/HICSS.2012.630>
- Rosas, S. R., & Ridings, J. W. (2017). The use of concept mapping in measurement development and evaluation: Application and future directions. *Evaluation and Program Planning*, 60, 265–276. <https://doi.org/10.1016/j.evalprogplan.2016.08.016>
- Salah, D. A., Paige, R., & Cairns, P. (2014). An evaluation template for expert review of maturity models. *Product-Focused Software Process Improvement: 15th International Conference, PROFES 2014, Helsinki, Finland, December 10-12, 2014. Proceedings 15*, 318–321. [https://doi.org/10.1007/978-3-319-13835-0\\_31](https://doi.org/10.1007/978-3-319-13835-0_31)
- Saputra, D. A., Handika, D., & Ruldeviyani, Y. (2018). Data Governance Maturity Model (DGM2) Assessment in Organization Transformation of Digital Telecommunication Company: Case Study of PT Telekomunikasi Indonesia. *2018 INTERNATIONAL CONFERENCE ON ADVANCED COMPUTER SCIENCE AND INFORMATION SYSTEMS (ICACISIS)*, 325–330. <https://doi.org/10.1109/ICACISIS.2018.8618255>
- Schumacher, A. (2015). *Development of a maturity model for assessing the industry 4.0 maturity of industrial enterprises* [PhD Thesis].
- Schumacher, A., Erol, S., & Sihn, W. (2016). A maturity model for assessing Industry 4.0 readiness and maturity of manufacturing enterprises. *Procedia Cirp*, 52, 161–166. <https://doi.org/10.1016/j.procir.2016.07.040>
- Soliman, J., & Kan, M. (2004). Grounded theory and NVivo: Wars and wins. *International Conference on Qualitative Research in IT & IT in Qualitative Research*.
- Soni, B., B., Bobby. (2023). Digital transformation. *KPMG Global*. Retrieved from <https://kpmg.com/xx/en/home/insights/2023/09/digital-transformation.html> (accessed 24 October 2023).
- Steinlechner, M., Schumacher, A., Fuchs, B., Reichsthaler, L., & Schlund, S. (2021). A maturity model to assess digital employee competencies in industrial enterprises. *Procedia CIRP*, 104, 1185–1190. <https://doi.org/10.1016/j.procir.2021.11.199>
- Sulaiman, H., Cob, Z. C., & Ali, N. (2015). Big Data Maturity Model for Malaysian Zakat Institutions to Embark on Big Data Initiatives. In J. Zain, M. Majid, M. Ameen, R. Arshah, Z. Azmi, & Z. Mustaffa (Eds.), *2015 4TH International Conference On Software Engineering and Computer Systems (ICSECS)* (pp. 61–66). Univ Malaysia Pahang, Fac Comp Syst & Software Engn; IEEE.
- Support. (2024). *MindManager*. Retrieved from <https://www.mindmanager.com/en/support/> (accessed 24 October 2023).
- Thabet, W., Lucas, J., & Srinivasan, S. (2022). Linking life cycle BIM data to a facility management system using Revit Dynamo. *Organization, Technology & Management in Construction*, 14(1), 2539–2558. <https://doi.org/10.2478/otmcj-2022-0001>
- Verdonck, M., Gailly, F., De Cesare, S., & Poels, G. (2015). Ontology-driven conceptual modeling: A systematic literature mapping and review. *Applied Ontology*, 10(3–4), 197–227. <https://doi.org/10.3233/AO-150154>
- Vesset, D., Girard, G., Feblowitz, J., Versace, M., Burghard, C., O'Brien, A., ... & Bond, S. (2015). IDC MaturityScape: big data and analytics 2.0. *IDC MaturityScape White Paper*.
- Xiang, J. (2022). Evaluation of Enterprise Accounting Data Management Based on Maturity Model. *Computational Intelligence and Neuroscience : CIN*, 2022. <https://doi.org/10.1155/2022/3891475>
- Zitoun, C., Belghith, O., Ferjaoui, S., & Gabouje, S. S. D. (2021). DMMM: Data Management Maturity Model. *Proceedings - 2021 International Conference on Advanced Enterprise Information System, AEIS 2021*, 33–39. <http://dx.doi.org/10.1109/AEIS53850.2021.00013>
- Zoubek, M., & Šimon, M. (2021). The Framework for Logistics 4.0 Maturity Model with a Specification for Internal Logistics. *MM Science Journal*, 2021(1), 4264–4274. [https://doi.org/10.17973/MMSJ.2021\\_03\\_2020073](https://doi.org/10.17973/MMSJ.2021_03_2020073)