Outpacing the Competition: A Design Principle Framework for Comparative Digital Maturity Models

Maximilian Breitruck[a]
Institute of Information Systems and Digital Business, University of St. Gallen, St. Gallen, Switzerland

Keywords: Digital Maturity Models, Benchmarking, Design Principles.

Abstract: Digital maturity models (DMMs) already have a long history of providing organizations with structured approaches for assessing and guiding their digital transformation initiatives. While descriptive and prescriptive DMMs have seen extensive development, comparatively few models focus on benchmarking digital maturity internally as well as externally across multiple organizations. Moreover, existing literature frequently highlights persistent shortcomings, including limited theoretical grounding, methodological inconsistencies, and inadequate empirical validation. This study addresses these gaps by synthesizing insights from a systematic literature review of 58 publications into a cohesive set of design principles for comparative DMMs. We differentiate between "usage design principles," which adapt established descriptive and prescriptive DMM components to comparative contexts, and newly formulated principles developed specifically to accommodate implicit data sources and support ongoing benchmarking. The resulting framework provides researchers and practitioners with a foundation for designing, evaluating, and selecting comparative DMMs that are more conceptually robust, methodologically sound, and empirically viable. Ultimately, this work aims to enhance the overall maturity and applicability of comparative DMMs in advancing organizational digital transformation.

1 INTRODUCTION

For decades, Digital Maturity Models (DMMs) have played a central role in digital transformation research within the discipline of Information Systems, leading to the development of numerous models. Nevertheless, recent publications on the subject consistently identify similar weaknesses in DMMs and propose areas for improvement (Thordsen et al., 2020; Thordsen & Bick, 2023). Recurring critiques include insufficient theoretical grounding, methodological inconsistencies, limited value creation, and the lack of a clear link between digital maturity and organizational performance—a connection often referenced as justification for advancing these models (Schallmo et al., 2021; Teichert, 2019). These critique points continue to be regarded as significant weaknesses in existing DMM frameworks. In an effort to address these issues, several publications have already proposed research agendas. Such agendas aim to clarify the specific challenges within DMM research and outline preliminary research questions that, if explored, could advance the field and contribute to the maturation of DMM research (Thordsen et al., 2020; Thordsen & Bick, 2023). To improve accessibility, research agendas can be simplified into two main directions essential for advancing DMMs: expanding the theoretical foundation and empirically validating the models—particularly the construct of digital maturity (DM) in relation to tertiary factors like performance and sustainability. The former focuses on how the creation of DMMs can be made more methodologically rigorous, which additional management theories can be applied, how the quality of DMMs can be assessed, and how general standards can be established. The empirical direction centers on determining whether and how DMMs generate real value in practice. This includes examining if higher digital maturity has statistically measurable impacts on factors such as organizational performance, stock prices, sustainability, customer satisfaction and access to external capital. Despite having fundamentally different goals, these two areas are

[a] https://orcid.org/0009-0000-1741-2042

Breitruck, M.
Outpacing the Competition: A Design Principle Framework for Comparative Digital Maturity Models.
DOI: 10.5220/0013427700003929

In *Proceedings of the 27th International Conference on Enterprise Information Systems (ICEIS 2025) - Volume 2*, pages 746-757
ISBN: 978-989-758-749-8; ISSN: 2184-4992

interrelated. Statistical significance in observed effects and high usability for model users are best achieved with a theoretically grounded, rigorously developed, and high-quality model. The theoretical contributions to DMM research thus indirectly influence the empirical dimension, as advancements in model quality enable more robust empirical investigations. A central aspect of the theoretical foundation of DMMs is seen in the development process. In existing research this is primarily associated with the Design Science Research (DSR) approach, widely regarded as the dominant method for creating DMMs. This design process, firmly established in the literature, provides detailed guidance on how to develop a DMM. However, as already noted by Pöppelbuß and Röglinger (2011), the components that are often overlooked are the form and function of these models. At the same time research repeatedly demonstrates that such comparative models are essential for conducting empirical studies on the relationship between maturity and other factors. For example, a statistical correlation analysis between maturity and organizational performance would require a maturity index tracked over time across multiple organizations. These types of scores are inherently developed in the creation of comparative models, as they form the basis for effective benchmarking. research question for this study is therefore: *What general design principles should comparative maturity models comply with to maximize their usefulness within their application domain and for their intended purpose?* To systematically address this question, the following chapters will present a systematic literature review aimed at synthesizing and structuring design principles that enhance the utility of comparative DMMs and digital maturity benchmarking approaches. These principles will be compiled into a structured framework, providing a foundation for both researchers and practitioners to more effectively design and evaluate comparative DMMs.

2 THEORETICAL BACKGROUND

2.1 Comparative Digital Maturity Models

When it comes to organizations the term maturity is defined as a "reflection of the appropriateness of its measurement and management practices in the context of its strategic objectives and response to environmental change," or more pragmatically as a "measure to evaluate the capabilities of an organization in regard to a certain discipline" (Thordsen et al., 2020, p. 360). Maturity models assess an organization's maturity level within a specific domain, such as processes or technologies in use. These models differentiate phases representing levels of maturity (Becker et al., 2010; Paulk et al., 1993). Applying maturity models involves a three-step process: assessing the current state (often via tools like questionnaires), defining a target state, and bridging the gap by planning subprojects and monitoring progress. DMMs evaluate and enhance digital capabilities, often considered equivalent to the degree of digital transformation. Unlike linear definitions imply, digital maturity is not a fixed endpoint; pathways to maturity depend on optimized factors and contingency considerations (Mettler & Ballester, 2021). Moreover, the required capabilities in digital transformation constantly evolve, making full maturity a temporary state until new advancements arise (Braojos et al., 2024; Rogers, 2023). Traditionally, this ongoing organizational development has only been addressed inefficiently by DMMs. These models require continuous application, with comparisons to competitors revealing an organization's relative digital maturity in practical terms (Minh & Thanh, 2022; Thordsen & Bick, 2023). The neglect of this aspect is however not due to the absence of a corresponding DMM type. In the foundational differentiation by Pöppelbuß and Röglinger (2011), the comparative DMM type is listed alongside descriptive and prescriptive types but has received little attention in research. Comparative DMMs are defined as models that "allow for internal or external benchmarking" (p. 5) Developing such models requires "sufficient historical data from a large number of assessment participants," (p.5) which explains their limited presence in existing research (Pöppelbuß & Röglinger, 2011). There is a limited amount of consulting surveys addressing the research gap by collecting maturity assessments over time (Rossmann, 2018). However, frequent assessments are impractical for continuous assessment. Textual datasets have emerged to overcome this limitation in academic research, often functioning as means in studies focussing on the correlation between digital maturity and economic or sustainability performance. For example, Hu et al. (2023) developed a Digital Transformation Index for China, generating digital maturity scores for all publicly listed Chinese companies across multiple periods. Based on a keyword lexicon from digital maturity literature, the index uses annual reports as the primary data source to identify digital capabilities, enabling relative rankings. This ranking was used to explore the relationship between corporate maturity mismatch and enterprise digital transformation. Although creating a comparative DMM was not the primary goal, it was implicitly achieved for empirical analysis. Similarly, Guo & Xu (2021) analyzed digital maturity in manufacturing firms using annual reports and a custom keyword library tailored to the sector, selecting 53 specific keywords. To address the limitation of annual reports being published yearly with limited digital transformation data, Ashouri et al. (2024) applied web scraping. This method analyzed annual reports and additional company information, identifying digital transformation efforts at product and organizational levels. Yamashiro & Mantovani (2021) also utilized scraping, focusing on external news websites to create a dataset for the Brazilian market. Axenbeck & Breithaupt (2022) combined these approaches, integrating corporate website data with news from Germany's largest outlets. Using a transfer learning model, the authors assessed the digital maturity of German companies. The literature on developing actual DMMs for benchmarking is more limited. Tutak & Brodny (2022) used a European Statistical Office dataset to compare maturity levels of countries based on local organizations. Their DMM analyzed eight ICT-related indicators. Warnecke et al. (2019) focused on Smart Cities, evaluating public sector organizations via a web platform and incorporating a prescriptive model. Being able to repeatedly use the model, highlighted changes over time and facilitated benchmarking as the user base grew. (Breitruck et al., 2024) expanded on these methodologies, using a significantly larger dataset of 1 billion pre-processed articles. Their approach explicitly focused on DMMs as benchmarking tools.

2.2 Design-Oriented Research

To better understand utility and identify principles that enhance utility in the design of comparative DMMs, a fundamental definition of the term is essential. In this research, DMMs and similar concepts for assessing digital maturity are regarded as artifacts which are created and validated through a design process. While comparative DMM studies do not exclusively rely on Design Science Research (DSR), its principles are observed in alternative methodologies. DSR defines artifacts broadly, encompassing design theories, principles, models, or guidelines, depending on the problem's maturity and application context (Baskerville et al., 2018). Notably, in DSR, the sequence of deriving design principles from requirements and using them to validate an artifact is often reversed. In practice, artifacts are first created and validated, afterwards, principles or theories are developed. Though counterintuitive, this approach aligns with DSR's iterative nature. Two core principles distinguish DSR: research-oriented design and design-oriented research. In practice, artifacts developed by consultants or industry focus on functionality and completion. By contrast, design-oriented research emphasizes knowledge generated during the design process, often resulting in experimental prototypes. This distinction applies to DMMs: consulting firms create detailed models without emphasizing gained knowledge or methodology. Academic publications, however, focus on both the artifact and insights from its creation. This contrast is reinforced by DSR's core principles, ensuring findings are methodologically sound and externally validated (Fallman, 2007). The two foundational principles of DSR, relevance and rigor, serve as pillars to ensure artifacts are both practically relevant and methodologically robust. To uphold these, numerous process models, guidelines, and principles have been developed to aid researchers in creating context-specific artifacts. In DMMs, publications provide guidance through design principles. Becker et al.'s (2009) model is widely recognized as a benchmark for DMM development and has been advanced by Mettler & Peter. For DMM content, Pöppelbuß & Röglinger's work is a standard reference. To identify design principles for comparative DMMs while adhering to DSR's usability and rigor criteria, Hevner's (2007) three-cycle model provides a useful framework. Beyond the design cycle, where the artifact is constructed, the model includes a relevance cycle and a rigor cycle. In the relevance cycle, an artifact is relevant if it improves a specific environment, determined by establishing requirements and evaluation criteria. The rigor cycle focuses on selecting and applying appropriate theories and methods while transparently communicating contributions to theory and methodology, including design products, processes, and project experiences (Hevner, 2007). DPs, applied in the rigor cycle, draw on pre-existing knowledge and ensure alignment with validated norms for artifact development. In the design phase, DPs also contribute to relevance. Through iterative construction and evaluation, artifacts are developed in alignment with DPs, ensuring they meet established standards and address the problem space. DPs can specify what users should achieve with the artifact, what features it should include, or both—defining features to support specific user activities (Gregor et al., 2020).

3 METHODOLOGY

The central objective of this research is to identify and organize existing design knowledge for developing comparative DMMs. To achieve this, relevant literature will be screened through a systematic literature review. The resulting knowledge will then be organized as nascent design theory, expressed as design principles, following the schema proposed by Gregor et al. (2020). These principles will be applied when evaluating existing DMMs to illustrate their use in validation and implications for the design cycle.

3.1 Systematic Literature Review

The following section outlines the literature search process. As established in previous research, a systematic approach to literature analysis was adopted following the methodologies of vom Brocke et al. (2015) and Webster & Watson (2002). Since the theorization and development of design knowledge, as previously mentioned, follow a counterintuitive process, literature on the topic of comparative DMMs will be identified and evaluated with respect to implicit and explicit design decisions. The search scope, 2011 to 2024, was set in line with previous literature focusing on DMMs, marking the time frame since the publication of the first DMM. Regarding the potential outlets for the publications reviewed, the traditional scope based on the IS Senior Scholar Basket and the four major IS conferences was expanded. The adjustment was made in response to the aforementioned need of additional literature on comparative DMMs. Consequently, all publications connected to digital maturity published in an outlet with recognition in the current VHB rating were included in the search. The rationale for this approach stem from the fact that many existing models were developed directly within specific application contexts, particularly in the finance domain. As previously mentioned, this is where the potential of comparative DMMs, even though not always labeled as such, is currently highly recognized, especially when conducting correlation analyses between DM and factors such as financial performance. Given that literature on comparative DMMs originates from both Information Systems and Finance disciplines, two separate search strings were developed. Search string 1 sought to identify explicitly labeled comparative DMMs, resulting in 409 papers. Search string 2 aimed to locate publications utilizing DMM-like constructs to analyze relationships between digital maturity and other variables, even if the DMM design was not the primary focus. This search yielded 668 publications. After removing duplicates and conducting an initial screening based on titles and abstracts, the combined sample was reduced to 58 publications including additional publications identified in a thorough forward backwar search.

Table 1: Search string overview.

No.	Search string
1	TITLE-ABS-KEY ((benchmarking) AND ((digital AND transformation) OR (digital AND maturity)))
2	TITLE (performance OR sustainability) AND ((digital AND maturity) OR (digital AND transformation)))

3.2 Design Principle Schema

The synthesized knowledge is systematically transferred into design principles using the schema by Gregor et al. (2020). This ensures clear guidance on how each principle can be applied in constructing and evaluating comparative DMMs. In DSR literature, design principles are nascent design theory, offering prescriptive knowledge on achieving specific goals. They describe how an artifact should be designed to enable users to perform tasks successfully. Gregor et al.'s (2020) approach uniquely incorporates users and stakeholders into the formulation of prescriptive design knowledge, addressing socio-technical systems in which the artifact operates. Gregor et al.'s (2020) schema, central to this study, identifies seven components, four of which are actor roles. Implementers use the principle to create an artifact, while users employ the artifact to achieve goals. Enactors support or oversee goal achievement, whilst theorizers abstract knowledge from applications without being part of the principle itself. The remaining components include context, mechanism, aim, and rationale. Context defines the boundary conditions and implementation settings for the artifact. Mechanism describes the processes required for users to achieve their goals. Aim specifies the intended objective, and rationale justifies why the mechanism is appropriate for achieving the aim. In summary, Gregor et al. (2020) express a design principle as: "DP Name: For Implementer I to achieve or allow for Aim A for User U in Context C, employ Mechanisms M1, M2, ... Mn involving Enactors E1, E2, ... En because of Rationale R" (p. 1633). This framework will guide the structuring of synthesized design knowledge into design principles in the following chapter.

4 PRINCIPLES FOR COMPARATIVE MATURITY MODELS

In the following section, the design principles formulated based on the synthesized insights from the literature are proposed. As shown in Figure 1, a fundamental distinction between two areas was made: Usage Design Principles (UP) and Design Principles. The term "Usage" in the first category was selected based on the findings of the literature review and existing DMM literature. Certain aspects of comparative DMMs can be adapted from other types of DMMs, such as descriptive and prescriptive models, as discussed by Pöppelbuss and Röglinger (2011). Consequently, these principles no longer need to address the fundamental design but focus on how other DMMs' existing components might need to be adjusted for use in comparative DMMs. This approach not only simplifies the construction and evaluation of comparative DMMs but also addresses the diversity of existing DMMs highlighted in the research. It aligns with the notion that models should not be developed solely for their own sake but only when a fundamental difference justifies their creation. By facilitating the recycling of existing models through the proposed UPs, this approach contributes to greater efficiency and alignment with the practical needs of DMM building (Thordsen et al., 2020; Thordsen & Bick, 2023).

UP1: The first UP cluster focuses on the foundation of DMMs, leveraging existing approaches to the fundamental construction of DMMs. Key aspects highlighted in the literature include the application domain (UP 1.1), the purpose of use (UP 1.2), as well as the design process and its associated validation (UP 1.3). Starting with the definition of the application domain, this qualifies as a usage principle (UP) because it is already established in the DMM literature as a fundamental design guideline. This includes identifying any prerequisites required to apply the model. A deviation from existing DMM types must be considered if the model is intended to work with larger, implicit datasets that may only be available for specific industries or publicly listed companies. Therefore determining whether the model is tailored to a specific industry, company size or is rather designed to be applied on a broader basis, independent of such factors, is crucial *(UP 1.1)* (Becker et al., 2009; Haryanti et al., 2023; Pöppelbuß & Röglinger, 2011; Thordsen et al., 2020; Thordsen & Bick, 2023).

Closely related to the application domain is the purpose of use, which presents specific considerations for comparative DMMs, as their primary purpose is to conduct a benchmarking of digital maturity levels. It is necessary to further specify whether this involves internal or external benchmarking, whether it tracks progress over time, or by contrast remains purely static. The benchmarking component is diverse, as the term is not always understood synonymously in the existing research *(UP 1.2)* (Becker et al., 2009; Pöppelbuß & Röglinger, 2011).

The same applies to the design process and the associated empirical validation of the model. Existing components of DMM development, based on the DSR methodology, must be slightly adjusted for comparative DMMs. Established processes for development (e.g., Becker et al., 2009; Mettler & Ballester, 2021) can certainly be utilized, but modifications are necessary to adapt them for the comparative DMM type. The benchmarking functionality is a key component of comparative DDMs, which, based on a calculation logic, enables the digital maturity of different organizations to be compared. This calculation significantly exceeds the complexity of traditional descriptive or prescriptive DMMs, particularly when large implicit datasets are used. As a result, clearer communication and documentation regarding the composition of these components during the design process are required n order for potential users to understand the basis of the benchmarking calculations. In this context, the extent to which the developed model has undergone empirical validation to demonstrate its validity must also be clearly communicated *(UP 1.3*) (Axenbeck & Breithaupt, 2022; Hu et al., 2023; Wu et al., 2024a).

UP2: The second Usage Design Principle cluster pertains to the structure of the model. Key aspects include the definition of the concept of digital maturity (UP 2.1), as well as digital maturation (UP 2.2). Starting with the central construct of DMMs—the definition of digital maturity—it becomes clear that fundamental design principles from descriptive and prescriptive DMMs can also be applied to comparative DMMs. For reasons of comparability, it is even advisable to use already validated constructs from existing DMMs as the foundation. The most common form of DMMs in existing research is structured as capability maturity models, where digital maturity is divided into dimensions such as technology, processes, and leadership, which are then supported by associated capabilities that organizations should ideally implement *(UP 2.1)*

(Becker et al., 2009; Becker et al., 2010; Mettler & Ballester, 2021).

The same applies to the definition of the various maturity levels connected to the maturity construct and the fundamental path through which maturity develops and evolves. Historically, the literature has understood maturity as a linear concept which fundamentally only increases. The most widespread approach, based on the capability maturity model,

defines a fixed number of maturity levels, each tied to specific capabilities. For example, Dimension X might have a capability corresponding to a low to more advanced maturity levels, culminating in the highest capability representing the most advanced maturity level. This approach may or may not be applicable depending on the final assessment method and the data basis used for the benchmarking built on the maturity construct. Models relying on implicit datasets, such as those scraped from the internet, often use more granular methods, assigning scores from 0 to 100 to indicate maturity levels. In such

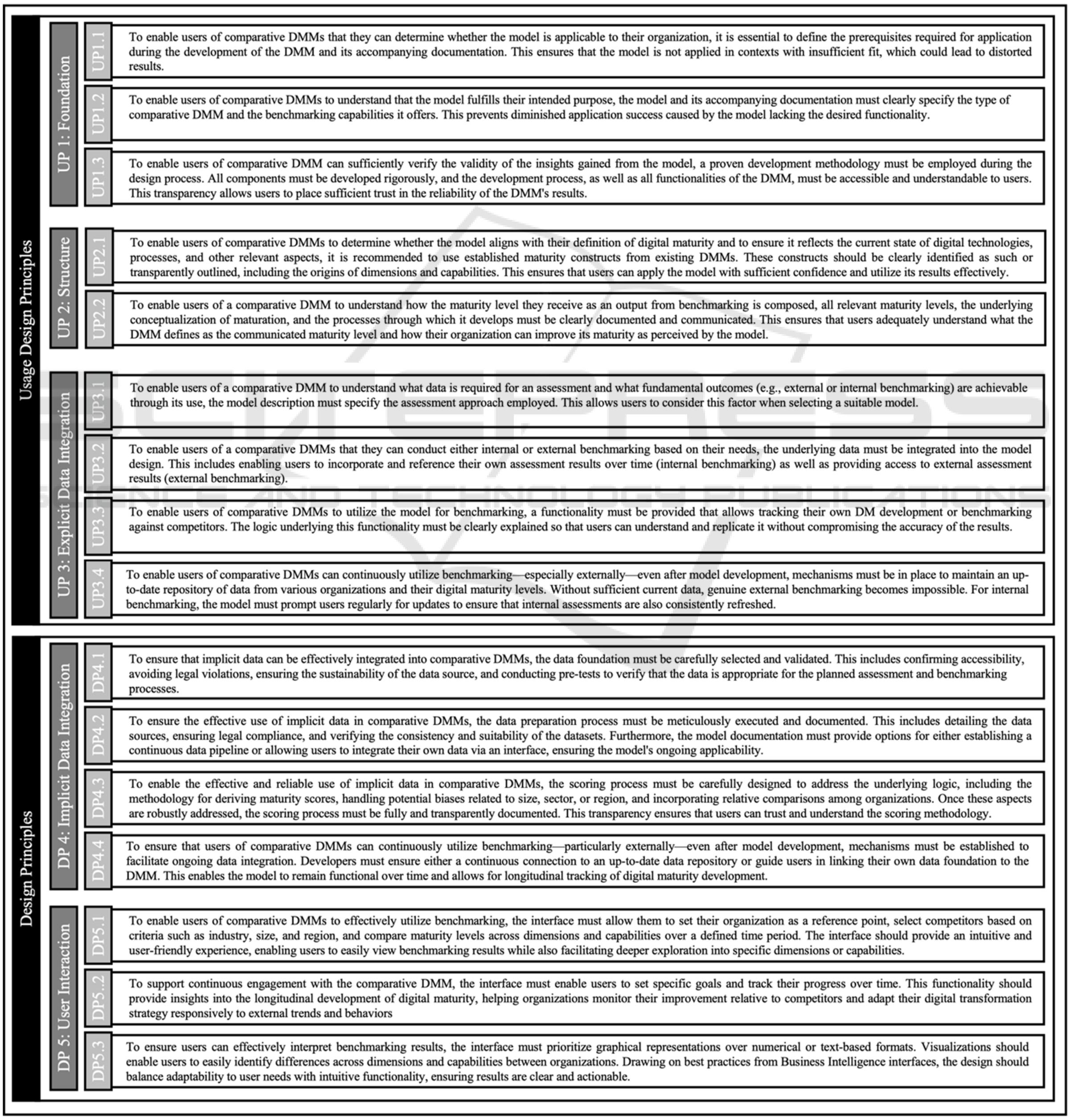

Figure 1: Principle framework for comparative DMMs.

cases, there may not be an explicit differentiation of 101 distinct maturity levels with their associated capabilities *(UP 2.2)* (Mettler & Ballester, 2021; Pöppelbuß & Röglinger, 2011; Soares et al., 2021; Teichert, 2019).

UP3: The third cluster of usage design principles addresses the actual functionality of the DMM. During the literature review process it became evident that in relation to comparative DMMs and adjacent digital maturity benchmarking approaches, two distinct methodologies regarding measurement techniques and data foundations are employed. The first approach, which is addressed in this cluster and already used in descriptive and prescriptive models, involves utilizing explicitly collected data, such as data from surveys (Berger et al., 2020; Berghaus & Back, 2017; Thordsen & Bick, 2023). The second approach, discussed in the next cluster, leverages implicit data gathered through methods such as web scraping. Starting with the assessment approach, it must be determined which type of data will be used. In alignment with existing DMMs, the use of explicitly collected data through questionnaires is a common practice. Based on the previously chosen structure, a questionnaire must be developed and sent to the organizations being surveyed to establish a consistent assessment infrastructure, enabling benchmarking. As discussed in the theoretical section, obtaining a sufficient number of responses is critical for external benchmarking, as this method relies on the comparison of organizations. This is not a requirement for internal benchmarking where only an organization's own periodic assessments are needed *(UP 3.1)* (Berghaus & Back, 2016; Pöppelbuß & Röglinger, 2011).

In the approach using explicit data, data access and preparation are relatively straightforward. For questionnaire data, a determination during the model design process whether internal benchmarking, external benchmarking, or both will be supported, is needed. If external benchmarking is used, mechanisms must be created to enable organizations to access assessment results from other organizations. This could be in form of a web-based tool with a repository of all assessment results, allowing for systematic and potentially anonymous data-sharing. This would enable organizations to compare their results with competitors, even in anonymized form *(UP 3.2*) (Kovačević et al., 2007; Wang et al., 2023; Warnecke et al., 2019).

A central functionality of DMMs, including the comparative type, is the scoring logic, which determines the specific DM level and provides applying organizations with insights into how mature they are based on the underlying DM model and data foundation. Here, as with descriptive and prescriptive DMMs, existing approaches can serve as a basis. If various specified maturity levels are used, these must be precisely defined and clearly distinguishable. In its simplest form, existing scoring systems can be used, wherein maturity levels across different DM dimensions are compared either with other organizations or with the organization's own past performance to conduct benchmarking. Additional factors, such as industry differences, company size, or specific transformation priorities, can also be incorporated to account for the relative nature of benchmarking. This ensures that benchmarking reflects not only static internal development but also DM development in relation to competitors. For instance, while an organization may achieve a sufficient maturity level in a self-evaluation, it may still lag significantly behind companies of the same size and in the same industry. It is crucial that the scoring logic aligns with the data foundation used and is transparently documented to ensure both accuracy and reproducibility *(UP 3.3)* (Barrane et al., 2021; Krstić et al., 2023; Wang et al., 2023; Warnecke et al., 2019).

Unlike descriptive or prescriptive models, benchmarking is not a static assessment. As already mentioned in UP 3.2 and UP 3.3, mechanisms must be established to ensure that the model and its underlying data are consistently updated, allowing for ongoing benchmarking over time. This continuous updating is essential to prevent the model from becoming obsolete shortly after its creation, where comparative data from various organizations might have been collected at a single point in time but not subsequently updated. Without this, benchmarking would only be possible against outdated data, reducing its relevance and accuracy *(UP 3.4)* (He & Chen, 2023; Warnecke et al., 2019; Wu et al., 2024b; Zhao et al., 2023).

DP 4: Building on the principles of UP 3, DP 4 provides concrete design principles that go beyond merely adapting existing guidelines for prescriptive and descriptive models. This is because implicit data types, such as scraped internet data, news data, annual reports, or other large textual datasets, have not been utilized in the previously mentioned model types. Therefore, new principles are required to integrate such data into comparative DMMs effectively. Starting with the assessment approach, as with explicit data, a clear definition of what and how it is being measured is needed. For implicit data, this specifically involves determining the data foundation to be used and how it will be leveraged for benchmarking. It is crucial to ensure that the data foundation meets certain basic criteria, such as accessibility, compliance with legal requirements, and the sustainability of the data source. Furthermore, a pre-test must confirm whether the selected data is fundamentally suitable for the intended assessment *(DP 4.1)* (Axenbeck & Breithaupt, 2022; Guo & Xu, 2021; Warnecke et al., 2019; Yamashiro & Mantovani, 2021).

As with explicit data, the assessment approach for implicit data also requires data extraction and preparation to enable the application of the corresponding calculation logic. However, this step is significantly more critical when using implicit data, as such data is not specifically collected or created for use in a DMM. This introduces a number of challenges, particularly regarding the extraction, consistency, and applicability of the data. For example, data from annual reports can be structured and extracted via financial data platforms like Bloomberg or Reuters Eikon and subsequently integrated into NLP pipelines to extract text from PDF documents. In contrast, web data, such as news articles or websites, requires building custom scraper scripts to extract structured data from the web. This data must then be processed into a uniform format, often requiring considerable effort. A clear documentation of the chosen approach is critical. This includes specifying which data was used, where it originated, ensuring its lawful usage, and verifying the consistency of the datasets. Additionally, to ensure the long-term usability of the model after its publication, the documentation must either facilitate the setup of a continuous data pipeline or allow users to integrate their own data via an interface *(DP 4.2)* (Axenbeck & Breithaupt, 2022; Liu, 2022; Yildirim et al., 2023).

The scoring logic for implicit data requires significantly more effort compared to explicit data, as individual maturity levels cannot be directly assessed during a survey but must instead be determined based on previously extracted datasets. The dominant approach in the literature involves creating keyword libraries that address DMM dimensions and capabilities. These libraries are then expanded using NLP tools, such as NLTK, to identify related terms. The enriched keyword set is subsequently applied to the extracted data to assess maturity. Scoring can follow traditional fixed maturity levels, identifying the maturity level of a capability for an organization based on the data. Alternatively, a keyword frequency approach can be used, where the word count of keywords within a capability cluster provides insights into its maturity. For instance, the frequency of a capability keyword appearing in the data for Organization X may correlate with higher maturity. Benchmarking introduces a relative comparison between organizations, requiring adjustments to static scoring methods. In the keyword frequency approach, scores are calculated based on how well an organization performs relative to its competitors. This adds complexity, as biases in the data—such as differences in organization size or regional characteristics—must be addressed. For example existing research suggests approaches like restricting comparisons to organizations within the same region or balancing data volumes to account for disparities in the amount of published news about organizations, even those of similar size. Therefore, it is critical to document the scoring process comprehensively and transparently, including the underlying logic, how biases related to size, sector, or region are handled, and how maturity scores are derived. This ensures users understand the methodology and can interpret the results with confidence *(DP 4.3)* (Axenbeck & Breithaupt, 2022; Guo & Xu, 2021; Tutak & Brodny, 2022; Warnecke et al., 2019; Yamashiro & Mantovani, 2021; Yildirim et al., 2023).

As with explicit data, the use of implicit data must also allow for continuous development, ensuring that the score evolves relative to the progress of competing organizations. As previously discussed, it is crucial that the data integration is either ensured by the developers or that users are guided in linking their own data foundation to the DMM. This approach ensures that the model can be used continuously after its creation and enables longitudinal measurement of digital maturity (DM) development. This allows organizations to observe how their DM evolves over time and progresses toward the benchmark, rather than merely providing a static snapshot of their current standing. This aspect is central to the development of comparative DMMs, enabling users to leverage the benchmarking functionality over time to track their progress in relation to digital maturity *(DP 4.4)* (Axenbeck & Breithaupt, 2022; Hu et al., 2023; Long et al., 2023; Warnecke et al., 2019; Wu et al., 2024b).

DP5: With regard to the last DP cluster, there is little connection to existing design knowledge or actual existing DMMs of the de- and prescriptive type. This is because the use of descriptive and prescriptive models has predominantly occurred without a user interface, relying solely on assessment tools composed of static materials such as textual documents. However, due to the dynamic nature of comparative DMMs, a dedicated interface between the model and the user in the form of an application interface is required. Whether this interface is web-based or a standalone locally executed application is secondary. What is more critical is that users can interact with the model effectively. To fulfill this purpose, the interface must address three key areas: functionality, visualization, goal setting & monitoring, together forming DP Cluster 5.

Starting with the core of the interface—the benchmarking functionality—it is essential that all elements outlined in the previous DPs are integrated and usable. The user must be able to set their organization as a reference point within the model, select competing companies based on industry, size, and region, and compare maturity levels broken down by dimensions and capabilities over a defined time period. The functionality should be designed to alleviate user's access to the benchmarking results and, if desired, delve deeper into specific areas of interest. This includes exploring individual dimensions or capabilities to gain more detailed insights. The usability and intuitiveness of this functionality are critical for ensuring that users can fully leverage the benchmarking capabilities of the DMM *(DP 5.1)* (Wang et al., 2023; Warnecke et al., 2019).

Closely related to this is the visual representation of the evaluation results. Existing research highlights that the way findings and scores are visualized significantly impacts how they are perceived and whether the DMM is considered useful. The core principle ensures that the visualization allows users to easily interpret the benchmarking results. The shift from numerical or text-based evaluations to graphical representations is particularly important, as it simplifies the identification of differences between organizations across dimensions and capabilities. Business Intelligence (BI) interfaces are often cited as a reference in existing research, as they are not only highly adaptable to user needs but also intuitive to use *(DP 5.2)* (Chuah & Wong, 2012; Chung et al., 2005).

Adjacent to the functionality of the interface is the option, frequently discussed in research, to go beyond traditional benchmarking by allowing users to set goals within the interface, similar to prescriptive models. This functionality enables organizations to track these goals over time, providing insights into their longitudinal development in DM. Such a feature encourages users to transition from one-time usage to continuous engagement with the tool. It therefore supports users in aligning their DT strategy responsively with environmental trends, competitor behavior, and other dynamics, fostering a more adaptive and proactive approach to DM improvement *(DP 5.3)* (Krstić et al., 2023; Wang et al., 2023; Warnecke et al., 2019).

5 DISCUSSION

DMMs offer organizations a scientifically grounded means to measure their maturity, advance it, and benchmark themselves both against their past performance and competing organizations. This enables them to adapt their DT strategy in response to internal progress and external changes. However, for such applications to be feasible, DMMs must incorporate essential components and functionalities that ensure their effective use by organizations. Over time, research has developed numerous approaches to designing DMMs and identifying the necessary components. However, the comparative DDM type has been largely overlooked, thus contributing to the limited development of such models. This is despite repeated calls from both researchers and practitioners for benchmarking and longitudinal approaches to measuring digital maturity, which comparative DMMs can provide.

To address this gap, a systematic literature analysis was conducted in order to identify and synthesize existing design and application knowledge from instantiated DMMs and related constructs. As described in Chapter 4, this knowledge was structured into design principles to make it accessible to both researchers and practitioners. Efforts, therefore, were made to adhere to rigorous scientific standards while ensuring the principles remain practically applicable. Researchers can use the derived principles as a foundation for the design and evaluation process of DMMs, while practitioners can rely on them to select/adapt a DMM suited to their specific organizational needs.

However, the chosen replicated research approach, namely the literature review, carries the significant limitation that the findings are based solely on existing published research, thereby lacking insights directly derived from practice. Additionally, it is important to note that the literature search heavily depends on the selected search strings. This dependency poses challenges, particularly in this field, where numerous related digital maturity constructs could be considered de facto instantiations of DMMs but are often referred to using different terminology. As a result, the search may have been incomplete, potentially overlooking theoretically relevant literature that is available.

To complete the final step of the literature review, the following two research fields are proposed as a future research agenda in this area.

Firstly, incorporating practitioner needs could be addressed through exploratory interviews to assess the use of the formulated principles or the application of comparative DMMs. Such an approach would provide deeper insights into the practical utilization of DMMs and the specific needs of practitioners.

Secondly, a systematic framework facilitating the adaptation of components from existing descriptive or prescriptive DMMs into comparative DMMs would be valuable. This would help reduce the creation of entirely new models and instead focus on leveraging proven components from existing models.

6 CONCLUSION

In summary, when designing comparative DMMs that enable organizations to compare their digital maturity either internally over time or externally against competing organizations, it is not always necessary to build entirely new models from scratch. Instead, targeted recycling and integration of existing DMMs can play a crucial role. By repurposing existing models or combining their elements with new features, it is possible to transform them into comparative DMMs or expand their functionality to meet the specific requirements of comparative benchmarking.

To align with this insight, the present paper does not aim to synthetically develop entirely new design principles for comparative DMMs. Instead, it focuses on promoting the integration of existing DMMs by formulating Usage Design Principles. These principles are intended to assist researchers and practitioners in understanding the connections among all three types of DMMs, while also avoiding unnecessary effort in creating new models. Suitable components, if already existent, can be recombined or repurposed rather than constructing new ones. For aspects of comparative DMMs where leveraging existing components or functionalities is not possible, additional design principles have been developed to provide a structured approach for incorporating these new elements. Together with the Usage Design Principles, they form a thorough framework that supports the design, evaluation, and selection process for comparative DMMs.

The principle framework, consisting of five clusters and a total of 16 principles, aims to provide value to both research and practice. By fostering the reuse and adaptation of existing models instead of unnecessarily creating new ones, researchers can be supported in streamlining the DMM design process. On the other hand, users can leverage previously established models for benchmarking or utilize the principles as a checklist when selecting a comparative DMM.

REFERENCES

Ashouri, S., Hajikhani, A., Suominen, A., Pukelis, L., & Cunningham, S. W. (2024). Measuring digitalization at scale using web scraped data. *Technological Forecasting and Social Change*, *207*, 123618. https://doi.org/https://doi.org/10.1016/j.techfore.2024.123618

Axenbeck, J., & Breithaupt, P. (2022). Measuring the digitalisation of firms–a novel text mining approach. *ZEW-Centre for European Economic Research Discussion Paper*, *22–065*.

Barrane, F. Z., Ndubisi, N. O., Kamble, S., Karuranga, G. E., & Poulin, D. (2021). Building trust in multi-stakeholder collaborations for new product development in the digital transformation era. *Benchmarking*, *28*(1), 205–228. https://doi.org/10.1108/BIJ-04-2020-0164

Baskerville, R., Baiyere, A., Gregor, S., Hevner, A., & Rossi, M. (2018). Design science research contributions: Finding a balance between artifact and theory. *Journal of the Association for Information Systems*, *19*(5), 3.

Becker, J., Knackstedt, R., & Pöppelbuß, J. (2009). Developing Maturity Models for IT Management. *Business & Information Systems Engineering 2009 1:3*, *1*(3), 213–222. https://doi.org/10.1007/S12599-009-0044-5

Becker, J., Niehaves, B., Poeppelbuss, J., & Simons, A. (2010). *Maturity models in IS research*.

Berger, S., Bitzer, M., Häckel, B., & Voit, C. (2020). *Approaching digital transformation-development of a multi-dimensional maturity model*. https://www.researchgate.net/profile/Stephan-Berger-2/publication/342262246_Approaching_Digital_Transformation_-_Development_of_a_multi-dimensional_Maturity_Model/links/5eeb4e1c458515814a676d61/Approaching-Digital-Transformation-Development-of-a-multi-dimensional-Maturity-Model.pdf

Berghaus, S., & Back, A. (2016). Gestaltungsbereiche der digitalen transformation von unternehmen: entwicklung eines reifegradmodells. *Die Unternehmung*, *70*(2), 98–123. https://www.nomos-elibrary.de/10.5771/0042-059X-2016-2-98.pdf

Berghaus, S., & Back, A. (2017). *Disentangling the Fuzzy Front End of Digital Transformation: Activities and Approaches*. http://aisel.aisnet.org/icis2017/PracticeOriented/Presentations/4

Braojos, J., Weritz, P., & Matute, J. (2024). Empowering organisational commitment through digital transformation capabilities: The role of digital leadership and a continuous learning environment. *Information Systems Journal*.

Breitruck, M., Back, A., & Pätzmann, L.-M. (2024). *Towards Assessing Digital Maturity: Utilizing Real-Time Public Data for Organizational Benchmarking*. https://www.alexandria.unisg.ch/entities/publication/79feae4b-b8dd-449f-a6d7-63aac08bd87b

Chuah, M.-H., & Wong, K.-L. (2012). Construct an enterprise business intelligence maturity model (EBI2M) using an integration approach: A conceptual framework. *Business Intelligence-Solution for Business Development*, 1–12.

Chung, W., Chen, H., & Nunamaker Jr, J. F. (2005). A visual framework for knowledge discovery on the Web: An empirical study of business intelligence exploration. *Journal of Management Information Systems*, *21*(4), 57–84.

Fallman, D. (2007). Why Research-Oriented Design Isn't Design-Oriented Research: On the Tensions Between Design and Research in an Implicit Design Discipline. *Knowledge, Technology & Policy*, *20*(3), 193–200. https://doi.org/10.1007/s12130-007-9022-8

Gregor, S., Chandra Kruse, L., & Seidel, S. (2020). Research perspectives: the anatomy of a design principle. *Journal of the Association for Information Systems*, *21*(6), 2.

Guo, L., & Xu, L. (2021). The Effects of Digital Transformation on Firm Performance: Evidence from China's Manufacturing Sector. *Sustainability 2021, Vol. 13, Page 12844*, *13*(22), 12844. https://doi.org/10.3390/SU132212844

Haryanti, T., Rakhmawati, N. A., & Subriadi, A. P. (2023). The Extended Digital Maturity Model. *Big Data and Cognitive Computing 2023, Vol. 7, Page 17*, *7*(1), 17. https://doi.org/10.3390/BDCC7010017

He, L. Y., & Chen, K. X. (2023). Digital transformation and carbon performance: evidence from firm-level data. *Environment, Development and Sustainability*. https://doi.org/10.1007/s10668-023-03143-x

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems*, *19*(2), 4.

Hu, Y., Che, D., Wu, F., & Chang, X. (2023). Corporate maturity mismatch and enterprise digital transformation: Evidence from China. *Finance Research Letters*, *53*, 103677. https://doi.org/https://doi.org/10.1016/j.frl.2023.103677

Kovačević, A., Kaune, S., Liebau, N., Steinmetz, R., & Mukherjee, P. (2007). Benchmarking Platform for Peer-to-Peer Systems (Benchmarking Plattform für Peer-to-Peer Systeme). *It-Information Technology*, *49*(5), 312–319.

Krstić, A., Rejman-Petrović, D., Nedeljković, I., & Mimović, P. (2023). Efficiency of the use of information and communication technologies as a determinant of the digital business transformation process. *Benchmarking*, *30*(10), 3860–3883. https://doi.org/10.1108/BIJ-07-2022-0439

Liu, Z. (2022). Impact of Digital Transformation of Engineering Enterprises on Enterprise Performance Based on Data Mining and Credible Bayesian Neural Network Model. *Genetics Research*, *2022*. https://doi.org/10.1155/2022/9403986

Long, H., Feng, G. F., & Chang, C. P. (2023). How does ESG performance promote corporate green innovation? *Economic Change and Restructuring*, *56*(4), 2889–2913. https://doi.org/10.1007/s10644-023-09536-2

Mettler, T., & Ballester, O. (2021). Maturity Models in Information Systems: A Review and Extension of Existing Guidelines. *ICIS*.

Minh, H. P., & Thanh, H. P. T. (2022). Comprehensive Review of Digital Maturity Model and ProposalforA Continuous Digital Transformation Process with Digital Maturity Model Integration. *IJCSNS*, *22*(1), 741.

Paulk, M. C., Curtis, B., Chrissis, M. B., & Weber, C. V. (1993). Capability maturity model, version 1.1. *IEEE Software*, *10*(4), 18–27.

Pöppelbuß, J., & Röglinger, M. (2011). *What makes a useful maturity model? A framework of general design principles for maturity models and its demonstration in business process management*. 28. https://aisel.aisnet.org/ecis2011/28/

Rogers, D. (2023). *The digital transformation roadmap: rebuild your organization for continuous change*. Columbia University Press.

Rossmann, A. (2018). *Digital maturity: Conceptualization and measurement model*.

Schallmo, D. R. A., Lang, K., Hasler, D., Ehmig-Klassen, K., & Williams, C. A. (2021). An approach for a digital maturity model for SMEs based on their requirements. In *Digitalization: Approaches, case studies, and tools for strategy, transformation and implementation* (pp. 87–101). Springer.

Soares, N., Monteiro, P., Duarte, F. J., & Machado, R. J. (2021). Extended Maturity Model for Digital Transformation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12952 LNCS*, 183–200. https://doi.org/10.1007/978-3-030-86973-1_13

Teichert, R. (2019). Digital Transformation Maturity: A Systematic Review of Literature. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, *67*(6), 1673–1687. https://doi.org/10.11118/actaun201967061673

Thordsen, T., & Bick, M. (2023). A decade of digital maturity models: much ado about nothing? *Information Systems and E-Business Management*, *21*(4), 947–976. https://doi.org/10.1007/s10257-023-00656-w

Thordsen, T., Murawski, M., & Bick, M. (2020). How to measure digitalization? A critical evaluation of digital maturity models. *Responsible Design, Implementation and Use of Information and Communication Technology: 19th IFIP WG 6.11 Conference on e-Business, e-Services, and e-Society, I3E 2020, Skukuza, South Africa, April 6–8, 2020, Proceedings, Part I 19*, 358–369.

Tutak, M., & Brodny, J. (2022). Business Digital Maturity in Europe and Its Implication for Open Innovation. *Journal of Open Innovation: Technology, Market, and*

Complexity, *8*(1). https://doi.org/10.3390/joitmc8010027

vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. *Communications of the Association for Information Systems*, *37*(1), 9. https://doi.org/10.17705/1CAIS.03709

Wang, C. N., Nguyen, T. D., Nhieu, N. L., & Hsueh, M. H. (2023). A Novel Psychological Decision-Making Approach for Healthcare Digital Transformation Benchmarking in ASEAN. *Applied Sciences (Switzerland)*, *13*(6). https://doi.org/10.3390/app13063711

Warnecke, D., Wittstock, R., & Teuteberg, F. (2019). Benchmarking of European smart cities–a maturity model and web-based self-assessment tool. *Sustainability Accounting, Management and Policy Journal*, *10*(4), 654–684.

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, *26*(2), xiii–xxiii. http://www.jstor.org/stable/4132319

Wu, X., Li, L., Liu, D., & Li, Q. (2024a). Technology empowerment: Digital transformation and enterprise ESG performance-Evidence from China's manufacturing sector. *PLoS ONE*, *19*(4 April). https://doi.org/10.1371/journal.pone.0302029

Wu, X., Li, L., Liu, D., & Li, Q. (2024b). Technology empowerment: Digital transformation and enterprise ESG performance-Evidence from China's manufacturing sector. *PLoS ONE*, *19*(4 April). https://doi.org/10.1371/journal.pone.0302029

Yamashiro, C. S., & Mantovani, D. (2021). Digital Transformation: digital maturity of the Brazilian companies. *RELCASI*, *13*(1), 2.

Yildirim, N., Gultekin, D., Hürses, C., & Akman, A. M. (2023). Exploring national digital transformation and Industry 4.0 policies through text mining: a comparative analysis including the Turkish case. *Journal of Science and Technology Policy Management*. https://doi.org/10.1108/JSTPM-07-2022-0107

Zhao, F., Meng, T., Wang, W., Alam, F., & Zhang, B. (2023). Digital Transformation and Firm Performance: Benefit from Letting Users Participate. *Journal of Global Information Management*, *31*(1). https://doi.org/10.4018/JGIM.322104