

Development of a Digital Curation Maturity Model and Indicators Including Weights for Measuring Digital Transformation Outcomes

Seonghun Kim*, Jinho Park**

ARTICLE INFO

Article history:

Received 27 February 2025

Revised 08 April 2025

Accepted 17 April 2025

Keywords:

Digital Curation,
Maturity Model,
Open Science,
Digital Transformation,
Confirmatory Factor Analysis

ABSTRACT

Information management institutions consider digital curation as a crucial area, and achieving digital transformation using the latest technologies is highly important. However, evaluation models for the maturity of digital curation and its continuous development have been scarce. Hence, this study developed a digital curation maturity model and indicators for measuring digital transformation outcomes. Weights were assigned to each indicator to allow for practical application during evaluations. This study identified 16 medium categories under 5 main categories, along with 40 subcategories. It also found 117 indicators and calculated weights for each main category and subcategory. The digital transformation maturity assessment model's most basic function is to assess the current state of the organization, providing a means for control and suggesting actions for the future. The findings can be used to conduct a systematic evaluation of digital curation maturity, contributing to the continuous development of information management.

1. Introduction

1.1 Necessity and purpose of the research

From an information management organization perspective, understanding and adopting new technological environments is an essential task, particularly when the technology is linked to the production and distribution of information resources. "Fourth Industrial Revolution" and "data" are said to be the keywords representing recent changes in the technological environment. The importance of technology, data, and information has further heightened because of the COVID-19 pandemic, acting as a catalyst especially in the expansion of information sharing and utilization to the general public.

To address these changes, various global information resource management and service organizations

* Affiliate Professor, iSchool - Library and Information Science, Sungkyunkwan University (SKKU) (godwmaw@skku.edu) (First Author)

** Assistant Professor, Department of Creative Humanities, Hansung University (jhp@hansung.ac.kr) (Corresponding Author)

International Journal of Knowledge Content Development & Technology, 15(3): 123-168, 2025.
<http://dx.doi.org/10.5865/IJKCT.2025.15.3.123>

are working toward digital transformation. In 2016, the International Institute for Management (IMD), in its annual World Digital Competitiveness Ranking report, introduced a specialized competitiveness index for digital transformation (IMD World Competitiveness Center, 2021). Furthermore, the German Engineering Federation developed IMPULS Industry 4.0 Readiness, which allows institutions or businesses to evaluate their preparedness for adapting to the Fourth Industrial Revolution. The Singapore government also introduced the Smart Industry Readiness Index (SIRI), which is based on a broad system of classifying processes, technologies, and organizations. The SIRI rates 16 detailed sub-categories on a six-point scale to assess digital maturity levels (Singapore Economic Development Board, 2020).

In South Korea, representative information resource management and service organizations that actively respond to digital transformation are the Korea Education and Research Information Service (KERIS), the National Library of Korea, and the Korea Institute of Science and Technology Information (KISTI). The KERIS places emphasis on the need to establish inclusive future education governance and proposes principles to adhere to in the digital transformation of future education (Korea Education and Research Information Service, 2021). The National Library of Korea formulated a three-year plan for digital services with a list of 15 detailed initiatives to address digital transformation (National Library of Korea, 2021).

Meanwhile, the KISTI not only worked at a theoretical level but also actively initiated business process reengineering (BPR) projects for digital transformation in 2021, seeking organizational workflow changes for digital transformation. Its implementation of BPR adopted the concept of a maturity model (National Science and Technology Data Center Content Curation Center, 2020), providing distinctive features. The maturity model is designed to evaluate an organization's capacity for continuous improvement; higher maturity levels correspond to lower probabilities of issues and greater adaptability to changes for quality improvement. Suitable methods for fostering medium- to long-term transformations include developing a maturity model, planning for changes in the current work environment, and evaluating performance based on outcome indicators. Using a maturity model allows for not only quantitative measurements but also qualitative assessments, making it even more versatile.

However, even if an institution aspiring for digital transformation analyzes and derives improvement measures for each of its departments' unit tasks, approaches to measure the outcomes of these improvement measures have been inadequate. Simply put, while plans are formulated, there is an absence of methods to assess outcomes.

Therefore, this study aimed to develop a digital curation maturity model and indicators for measuring digital transformation outcomes using a maturity model that assesses an institution's continuous development. Additionally, each indicator was assigned weights to allow for practical application during evaluations. The concept and scope of digital transformation can vary depending on the application. This study focused on open science-based digital transformation, specifically the continuous development and dissemination of scientific research outcomes.

1.2 Research scope and methods

To achieve this research goal, one must first conduct various case studies and, based on them,

construct scales and indicators. Key concepts in the construction of scales and indicators focus on digital transformation and data. With regard to data, considering actual tasks in the field, it is essential to include data quality, research data, open data, and even artificial intelligence (AI) learning data. Research methods for developing the digital curation maturity model and measurement indicators can be broadly categorized into two: The first involves a preliminary model construction based on different case studies and their results, considering the research objectives. The second entails model validation targeting experts, including relevant institutions and academia. These are summarized in Table 1.

Table 1. Study scope and method summary

Scope and Procedure		Method
Preliminary Model Configuration	Case study	Case of digital transformation, data quality management (DQM) model, domestic/international research and cases related to the DQM model
	Indicator derivation	Derivation of measurement indicators for each case study target
	Indicator summary	Organizing each indicator by dividing it into scales and measurement indicators (removal of duplicates, etc.)
Model Validation	Preliminary model configuration	Constructing the final preliminary model after refining scales and indicators
	Survey	Online survey
	Data cleaning	Cleaning survey results data
	Frequency analysis	Frequency analysis to analyze the descriptive statistics of collected data
	Reliability verification	Using Cronbach's α to verify whether the items related to the construct concept of the measurement tool (evaluation model) have internal consistency
Confirmatory factor analysis	Performing confirmatory factor analysis based on factor loading, AVE, and concept reliability to confirm the convergent and discriminant validity of the measurement tool (evaluation model)	

The preliminary model was constructed by eliminating redundancy and integrating scales and indicators suggested by different case studies that align with the research objectives. The model developed through the investigation was evaluated through user surveys. The surveys sought to determine the appropriateness of the model's measurement indicators; for this purpose, confirmatory factor analysis and reliability verification were performed.

Additionally, to derive weights for the measurement indicators of the digital curation maturity model developed in this study, the analytic hierarchy process (AHP) technique was employed. Proposed by Thomas L. Saaty in 1980, AHP is a method for finding solutions to complex decision-making problems. Saaty (1980) suggested a technique for decomposing various options in complex decision-making problems into components, determining these components' relative priorities, and deciding on the final priorities. Since then, the AHP has been used for decision-making in different fields. The general AHP procedure follows the steps outlined by Saaty (1980), summarized as follows:

- Step 1: Decompose the decision-making issue into a hierarchical structure.
- Step 2: Create matrices representing the relations between each layer and its sublayers.
- Step 3: Assign relative weights to each relation.
- Step 4: Evaluate the relative scores for each alternative based on how well they satisfy the criteria.
- Step 5: Aggregate the computed scores to comprehensively estimate the value of each alternative.

Because this study focused on deriving weights for existing models from research, it skipped step 1 and proceeded with the steps outlined in Table 2:

Table 2. Analytic hierarchy process application procedures and methods

Procedure	Method
Conducting a survey	<ul style="list-style-type: none"> • Selecting the survey target • Evaluating the relative importance between elements using a Likert scale
Creating a pairwise comparison matrix	<ul style="list-style-type: none"> • Creating a pairwise comparison matrix to compare relative importance between elements
Performing a consistency test	<ul style="list-style-type: none"> • Checking matrix consistency to ensure reliability of comparison (Saaty, 1980) • Calculating the consistency ratio (CR); if the CR is 0.1 or less, consistency is judged to be sufficient <ul style="list-style-type: none"> - $CR = CI / RI$ • Consistency index (CI) measurement <ul style="list-style-type: none"> - $CI = (\lambda_{max} - n) / (n - 1)$ - λ_{max} is the maximum eigenvalue of the matrix, n is the matrix size, and RI is the random CI.
Presenting the weighted model	<ul style="list-style-type: none"> • At the model level, propose a method for applying weights between indicators

2. Literature analysis

Studies on the evaluation of digital curation maturity have been scant, but those aiming to create evaluation indicators for similar concepts can be found in various fields. The following sections summarize these studies according to their fields.

2.1 Digital transformation and maturity assessment

Digital transformation involves a complete change in an organization's working methods, organizational culture, and more based on digital technologies through data utilization (Park & Cho, 2021). Digital transformation transcends digitization, which focuses on the conversion of analog data into digital data, and digitalization, which involves the use of information technology (IT) in business operations and processes based on digital data. Digital transformation represents a more advanced concept that revolutionizes an organization's overall culture, working methods, and thinking processes around digital technologies.

In the digital transformation context, digital maturity is crucial when evaluating an organization's adaptation or readiness level for the digital business environment. The concept of digital maturity has evolved from its original role in the field of information systems and software development, which was the assessment of an organization's holistic management capabilities affecting quality. Recent studies have actively examined digital transformation maturity evaluation models, and the term is now used to signify an organization's systematic preparation for consistently adapting to digital change (Heo & Cheon, 2021).

Therefore, existing research must be reviewed to achieve the main goal of this study, which is to construct a model for evaluating the level of digital transformation maturity in the context of content curation systems. Such analysis will help collect foundational information for the assessment framework and extract insights to organize considerations for model development. This study initially examined recent maturity assessment models related to digital transformation and proceeded to investigate general quality assessment models, service quality assessment models, process quality assessment models, and data-centric quality assessment models spanning open data, research data, and AI data.

2.2 Digital transformation assessment model

Research perspectives on the diagnosis and measurement of digital transformation maturity indicators can be categorized into (1) macroscopic, top-down and (2) microscopic, bottom-up. Studies adopting the top-down perspective are primarily conducted at the national level, defining industries associated with the digital economy and investigating metrics such as sales, employment, and research and development investments of these industries. It also involves the evaluation of broad indicators such as a country's overall digital accessibility, technological and human resource capabilities, institutional regulations, and social trust. Conversely, studies adopting the bottom-up perspective include cases where institutions (or companies) develop models for assessing digital transformation at an organizational or private unit level.

The IMD publishes its World Digital Competitiveness Ranking, focusing on major categories such as knowledge, technology, and future readiness, with a central framework consisting of 52 evaluation criteria (IMD World Competitiveness Center, 2021). Since 2002, the World Economic Forum has been publishing its Network Readiness Index (NRI), which evaluates the digital capacities of countries and consists of 60 detailed indicators classified under technology, people, governance, and impact. The Organization for Economic Co-operation and Development (OECD) also introduced a comprehensive digital policy framework for assessing national-level digital development status using 33 digital transformation measurement indicators categorized into seven areas: access, use, innovation, jobs, society, trust, and market openness. The SIRI, which was developed by the Singapore government, is based on the IMPULS Foundation's Industry 4.0 Readiness in Germany and evaluates digital maturity level through 16 detailed classification items under process, technology, and organization categories (Singapore Economic Development Board, 2020).

One notable model in the realm of evaluating institutional (or corporate) and private-unit-level digital transformation is the IMPULS Industry 4.0 Readiness developed by the German Engineering Federation (IMPULS, n.d.). This model rates institutions' or companies' readiness in adapting to

the Fourth Industrial Revolution. It has six categories: strategy and organization, smart factories, smart operations, smart manufacturing, data-driven services, and employees. Its 28 questions allow readiness levels to be determined across six stages. Gartner, a prominent IT research company in the United States, developed a public sector-focused digital transformation assessment tool that contains seven key indicators: vision and strategy, service delivery and quality, organization, organizational readiness, digital projects and investments, the CIO's role, and data and analytics. The World Bank's open data quality measurement model focuses on evaluating leadership, open data ecosystems, policy and legal frameworks, organizational responsibility structures within the government, government data, finance, national technology infrastructure, and citizen engagement.

In Korea, studies have examined the development of digital maturity models for digital transformation. The Korea Institute of Public Administration (2021) developed a digital transformation index model to measure digital transformation levels in the public sector, which consists of connectivity, automation, virtualization, and data-based indices. According to Heo and Cheon (2021), the digital maturity model comprises four dimensions: technological readiness, strategic readiness, organizational culture, and human resource readiness. Hong, Choi and Kim (2019) have constructed a digital transformation capability assessment model that is suitable for the domestic context, rating 32 detailed measurement indicators focusing on technological and organizational capabilities.

2.3 Data quality evaluation model

ISO 8000 defines data quality as the value of information assets that enhance business efficiency and support strategic decision-making by providing suitable and accurate data promptly, securely, and consistently (International Organization for Standardization [ISO], 2016, 2022). An organization's data quality management is a crucial aspect linked to its overall value. Therefore, before configuring indicators for the digital transformation maturity model, we aimed to construct a preliminary model based on a review of data quality measurement models.

ISO/IEC 9126 measures data quality through factors such as functionality, reliability, usability, efficiency, maintainability, and portability. ISO/IEC 25012 further advances these elements by suggesting 15 quality measurement factors: accuracy, completeness, consistency, reliability, currency, accessibility, compliance, confidentiality, efficiency, precision, traceability, understandability, usefulness, portability, and recoverability (International Organization for Standardization [ISO], 2001, 2008).

As a case for measuring open data and public data quality, the National Information Society Agency of Korea (NIA) uses seven indicators: readiness, completeness, consistency, accuracy, security, timeliness, and usefulness (NIA, 2018). Tim Berners-Lee's five-star open data is also a representative quality measurement metric. The Research Data Alliance's FAIR Data Maturity Model contains indicators for searchability, accessibility, interoperability, and reusability to set common evaluation criteria for research data (RDA FAIR Data Maturity Model WG, 2020).

With the increasing interest in AI data, many studies have investigated quality management requirements for such data. According to the Telecommunications Technology Association, quality measurement indicators may include diversity, comprehensiveness, volatility, reliability of sources, factuality, standard compliance, statistical sufficiency, statistical uniformity, suitability, and label accuracy

(Telecommunications Technology Association, 2021).

The NIA, in its AI Data Quality Management Guidelines, incorporated the opinions of various stakeholders to measure AI data quality through 10 indicators: readiness, completeness, usefulness, standard compliance, statistical diversity, semantic accuracy, syntactic accuracy, algorithmic adequacy, and validity (NIA, 2021; 2022). Shin (2021) proposed criteria for verifying AI training data in terms of diversity, syntactic accuracy, semantic accuracy, and validity. Additionally, according to Kim and Lim (2020), quality management items for AI training data include diversity, reliability, fairness, sufficiency, uniformity, factuality, suitability of annotation for functional purposes, clarity of object classification, comprehensiveness of annotation attribute information, and effectiveness of learning.

Research on data quality management systems or processes assumes that an organization's data quality management occurs not as a singular act at a specific point in time but as an integral part of the overall process. According to ISO 9001, which focuses on quality maintenance and assessment, a continuous plan-do-check-act (PDCA) cycle allows for ongoing business improvement and quality management. Building upon ISO 9001, ISO 8000-61 presents 20 quality management processes based on the cyclic structure of quality planning, quality control, quality assurance, and continuous quality improvement.

Capability maturity model integration (CMMI) is a model that conducts holistic assessments in process management, project management, engineering, and support. The Korea Institute of Information and Communication Technology Promotion developed the PCL quality management maturity model with reference to the plan-build-operate-utilize cycle and proposed the ACL maturity level assessment model to assess the capability levels of activities that constitute processes in subsequent research. Additionally, guidelines were provided to select suitable metrics and apply them during the life cycle of AI training data, encompassing planning, data acquisition, data refinement, data labeling, and data training processes.

In reviewing digital transformation maturity models at the national or institutional (private) level, quality management systems or models from a data perspective, and those in general software or service domains, organizations measure maturity levels by crucially utilizing evaluation factors associated with technical aspects, human resources, and governance. National-level evaluations address the societal and economic impact of digital transformation and highlight the link between digital maturity and a country's social and economic competitiveness. Moreover, existing maturity assessment models lack a dedicated provision for a thorough evaluation of the 'data' itself, which is a core management focus for organizations. Therefore, this study addresses the limitations of existing models by considering the addition of evaluation factors for the data itself and factors that assess institutional-level societal impact.

3. Preliminary model configuration for digital curation maturity assessment

The previous section conducted a comprehensive review that encompasses literature up to the digital maturity stage, including digital transformation evaluation models, data quality measurement

models, and cases of quality management process models. This analysis revealed different measurement items and scales as well as instances in which different names were used for the same meaning. Therefore, before constructing the preliminary model, this study reorganized the reviewed cases to differentiate scales and indicators and selected scales and indicators to be used in the model by refining and integrating duplicate elements. Table 3 outlines this process.

Table 3. Preliminary model derivation process and method

Method	How to do it
Assigning document identifiers	• Assigning identifiers to each case investigation target
Setting criteria	• Setting criteria to organize research and case study elements
Organizing elements according to criteria	• Organizing research and case study elements according to criteria
Integrating and refining redundancies	• Integrating the same scales and indicators
Constructing the final preliminary model	• Constructing the indicators and scales of the final preliminary model

3.1 Assignment of literature identification numbers

To organize the extensive results of the case studies, the first task was the assignment of identification numbers to each piece of literature. This was performed not only to organize but also to facilitate source verification when constructing elements such as scales or indicators in the future. This task entailed a distinction between cases of digital transformation, data quality measurement, data quality management stages, and maturity stages.

Table 4. Assignment of literature identification numbers for each case study subject for constructing a preliminary model

Category	Identification number	Literature and models to be investigated
Case of digital transformation	1-A	Digital transformation competency indicator (Hong, Choi, & Kim, 2019)
	1-B	Public digital level diagnosis model (Korea Institute of Public Administration, 2021)
	1-C	IMD Digital Competitiveness Ranking 2021 (IMD World Competitiveness Center, 2021)
	1-D	Network Readiness Index 2021 (Portulans Institute, 2021)
	1-E	Measuring the Digital Transformation (OECD, 2019)
	1-F	Digital Government Urgency, Readiness, and Maturity Assessment (Gartner, n.d.)
	1-G	The Readiness Measurement Model (IMPULS, n.d.)
	1-H	Digital Maturity Model for Digital Transformation (Heo & Cheon, 2021)
	1-I	The Smart Industry Readiness Index 2020 (Singapore Economic Development Board, 2020)

Category	Identification number	Literature and models to be investigated
Data quality measurement model	2-A	FAIR Data Maturity Model: Specification and Guidelines (RDA FAIR Data Maturity Model WG, 2020)
	2-B	OpenAIRE Metadata Quality Challenges (Principe et al., 2019)
	2-C	Springer Nature (Stuart et al., 2018)
	2-D	DATAONE-Article Data Center MetaDIG (DataOne, n.d.)
	2-E	Data quality control requirements for supervised learning (TTAK.KO-10.1339)
	2-F	Data quality management guidelines for artificial intelligence learning V1.0 (NIA, 2021)
	2-G	Data quality management guidelines for artificial intelligence learning V2.0 (NIA, 2021)
	2-H	Data quality verification method for artificial intelligence learning (Shin, 2021)
	2-I	Quality verification item for artificial intelligence learning data (Kim & Lim, 2020)
	2-J	ISO/IEC 9126 (https://en.wikipedia.org/wiki/ISO/IEC_9126)
	2-K	ISO/IEC 25012 (Jeong, 2007)
	2-L	Five-Star Open Data (http://5stardata.info/ko/)
	2-M	Public data quality management (NIA, 2018)
	2-N	Public big data platform performance evaluation model (Lee, Park & Ryu, 2020)
	2-O	SERVQUAL (Parasuraman & Berry, 1988)
Data quality management	3-A	Quality management systems — Requirement (ISO 9001)
	3-B	Public data quality management level evaluation model (PCL) (Kim, Lee, & Kim, 2015)
	3-C	Public data quality management maturity level evaluation model based on activity ability level (ACL) (Kim, Lee & Lee, 2017)
	3-D	NIA - Quality Management System (NIA, 2018)
	3-E	Data quality control requirements for supervised learning (TTAK.KO-10.1339)
	3-F	Data quality management guidelines for artificial intelligence learning v2.0 (NIA, 2022)
	3-G	ISO 8000-150 (https://itwiki.kr/w/ISO_8000-150)
	3-H	ISO 8000-61 Data Quality Management: Process Reference Model (https://www.dpadvantage.co.uk/2020/02/05/iso-8000-61-the-data-quality-management-standard/)
	3-I	CMMI (Lanin, 2008)

3.2 Criteria setting

The second stage involves setting criteria for designing the preliminary model. In other studies, indicators mostly followed a three-tier structure consisting of major, sub, and minor levels with diverse scopes and depths. To address this, criteria were established to merge similar items and perform consistent mapping. This was based on IMPULS Industry 4.0 from Germany, which is the most referenced model among studies on digital transformation maturity levels. The measurement

categories presented in this model include (1) strategy and organization, (2) smart factory, (3) smart operations, (4) smart products, (5) data-based services, and (6) personnel. In the case of strategy and organization, the concept of strategy for digital transformation and the organization's strategy were distinguished because other models often separate them and define them under different classification criteria. Specifically, organizations were frequently addressed separately, often in conjunction with personnel and other aspects. Therefore, the organization was a distinct category within the strategy and organization component.

With regard to smart factory, smart operations, smart products, and data-based services, the content included in other indicators were technology related and therefore integrated and changed into the technology category.

- Strategy and organization → Strategy
- Smart factory, smart operations, smart products, data-based services → Technology
- Personnel, strategy and organization → Organization (personnel)

Through this process, the major categories were organized into strategy, technology, and organization (personnel). An additional major category involved the data items under management, which were the main focus of this study.

3.3 Element organization according to criteria

Based on the major categories established in the previous stage—strategy, technology, organization (personnel), and data—the elements derived from each case were mapped and grouped into middle categories and subcategories. In this grouping process, the names of elements, along with their definitions, indicators, and measurement methods, were examined. Similarities were used as criteria to map and group elements that were identical or similar. A merging process was also undertaken for elements with nearly identical names and definitions.

Finally, elements outside the four major categories (e.g., economic influence in the NRI, contribution to quality of life) were classified as miscellaneous items, resulting in a total of five main categories: strategy, technology, data, organization (personnel), and miscellaneous.

3.4 Duplicate and integration refinement

Duplicates were removed and elements were integrated and refined according to the selected main categories. Because the previous steps involved grouping elements based on their names, descriptions, and indicators, this stage focused on verifying whether mapping and refinement had been appropriate.

Specifically, the task involved integrating or creating new categories for main, middle, and subclassifications. “(Social) impact” was ultimately modified to be included in the miscellaneous category. While the concept of (social) impact may not be easily visible in a microlevel assessment of organizational maturity, it is a significant factor at the macro level, especially in measuring

the extent of national-level digital transformation. (Social) impact measures the degree to which a country’s digital maturity level contributes to the lives of its citizens. While it might be challenging to observe in a micro-level assessment of organizational maturity, considering the economic, educational, and environmental impact on individuals, organizations, and society, based on the core “data” handled by the organization, it was deemed an important evaluation factor in measuring the organization’s digital transformation maturity.

Through the aforementioned process, indicators were mapped to construct the preliminary model. Table 4 shows that the focus was on the main categories of technology, data, strategy, organization (personnel), and (social) impact. Middle classification and subclassification were then conducted. To ensure the traceability of the content of all indicators, each indicator was assigned a source identification number based on its minor classification, and the final preliminary model was established.

Table 5. Classification system and indicator composition of preliminary model mapping

Main category	Middle category	Subcategory	Source of indicators
Technology	Research and development	R&D investment	1-A, 1-C
		Technology development capabilities	1-A, 1-I
		Innovation capability	1-A, 1-F, 1-H
	IT infrastructure	Accessibility	1-C, 1-D, 1-E, 1-I
		Security	1-D, 1-M
		Work utilization	1-A, 1-B, 1-D, 1-G, 1-F, 1-H, 1-I
		Network capabilities	1-A, 1-B, 1-H, 1-I
	Regulatory environment	Understanding the regulatory environment	1-D, 1-C
		Regulatory application and compliance	1-D, 1-C
Data	Data quality	Diversity	2-H, 2-E
		Compatibility	2-B, 2-J, 2-H, 2-E, 2-F, 2-G, 2-I, 2-M, 2-K, 2-N
		Usability	2-K, 2-A, 2-C, 2-M
		Timeliness	2-K, 2-E, 2-N, 2-J
		Interoperability	2-J, 2-A
		Security	2-J, 2-M
		Maintainability	2-J
	Data management process	Readiness	2-F, 2-M, 2-G 3-B, 3-C, 3-D, 3-G, 3-I
		Completeness	2-F, 2-G 3-E, 3-F, 3-I, 3-C
		Usefulness	2-F, 2-G 3-A, 3-E, 3-I, 3-H
Strategy	Organizational-level strategy	Vision and goal	1-A, 1-F
		Policies and processes	1-A, 1-F, 1-H, 1-I
		Process innovation	1-I

Main category	Middle category	Subcategory	Source of indicators	
Organization (personnel)	Sector-specific strategies	Policies and processes	1-A, 1-F, 1-G, 1-H, 1-I	
		Talent acquisition strategy	1-A, 1-C	
		Commercialization strategy	1-A, 1-H	
		R&D strategy	1-C, 1-D, 1-E	
		Service strategy	1-H, 1-F	
	Organization	Forming a dedicated organization	Organizational personnel composition	1-A
			Organizational skill	1-G
			Organizational connection	1-F
			Organizational connection	1-A
	Personal competency	Work initiative	Work initiative	1-G, 1-H
Work resilience			1-I, 1-H	
Readiness for change			1-H	
Technical management skills			1-A, 1-C, 1-D, 1-H	
Leadership competencies	Leadership system	Leadership system	1-A, 1-I	
		Executive CIO role	1-F	
Operation and management	Human resources management	Human resources management	1-A	
		Talent training	1-C, 1-E	
(Social) impact	Contribution to bridging the digital gap	Economic effect	1-E	
		Economic effect	1-D	
		Educational effect	1-D	
		Degree of data openness	1-I	
		Overall satisfaction	1-I	

3.5 Final configuration of the preliminary model

After reviewing the cases, setting the criteria, selecting elements, modifying and merging indicators, and undergoing the refinement process, the final preliminary model for measuring digital transformation maturity is created, shown in Figure 1. The evaluation scale consists of 5 main categories, 16 middle categories, and 40 subcategories. Each subcategory contains 37 evaluation indicators for technology, 45 for data, 18 for strategy, 36 for organization (personnel), and 14 for (social) impact.

Measure			Indicators
Major categories (5)	Medium category (16)	Subcategories (40)	Evaluation indicators corresponding to subcategories
<ul style="list-style-type: none"> • Technology • Data • Strategy • Organization(personnel) • (social)Impact 	<ul style="list-style-type: none"> • Technology(3) • Data(2) • Strategy(2) • Organization(personnel)(4) • (social)Impact (5) 	<ul style="list-style-type: none"> • Technology (9) • Data(10) • Strategy(8) • Organization(personnel)(13) • (social)Impact (0) 	<ul style="list-style-type: none"> • Technology (37) • Data (45) • Strategy (18) • Organization(personnel) (36) • (social)Impact (14)

Fig. 1. Configuration of measures and indicators of the preliminary model

The classification system for the preliminary model, including its main and middle categories along with definitions for the former, is illustrated in Figure 2.



Fig. 2. Model validation and final model derivation

4. Model verification and final model derivation

With regard to the preliminary digital maturity evaluation model derived in this study, it is crucial to confirm whether the conceptual constructs and measurement indicators accurately represent the concepts. To this end, this study performed confirmatory factor analysis to assess convergent and discriminant validity. Convergent validity assumes that if multiple measurement indicators are used to measure a single conceptual construct, various measurement indicators should be highly correlated (Noh, 2019). Meanwhile, discriminant validity assumes that if different conceptual constructs are measured through multiple measurement indicators, their correlation should be low. This validity

verification was conducted in three stages:

- Step 1: Critical ratio (CR) values were confirmed based on unstandardized λ values, ensuring that they are above 1.96 ($p < 0.05$).
- Step 2: Convergent validity was verified using the three criteria below, and measurement indicators that did not exceed the standardized λ value criteria were removed for further validation:
 - Standardized λ values exceeding 0.7: This study used a threshold of 0.7 as it is required in some specialized areas, while general social science research often uses 0.5 as a criterion (Noh, 2019; Yu, 2012).
 - Checking if average variance extracted (AVE) values are above 0.5.
 - Confirming that composite reliability (CR) values are above 0.7.
- Step 3: Discriminant validity was verified through two processes:
 - Checking if AVE values are greater than the square of the correlation coefficient.
 - Confirming that there is no “1” within the (correlation coefficient ± 2 *standard error) range.

However, in step 3, even though conceptual constructs were deemed independent in other studies, in cases where they significantly influence each other, especially with a high correlation, instead of excluding or integrating conceptual constructs, it is more appropriate to present studies that used them as independent concepts. Decisions were made after a careful consideration of their relevance in this study (Yu, 2012).

The confirmatory factor analysis survey was conducted online using Google Forms and SurveyMonkey from August 16 to September 13, 2022, targeting individuals who performed data management tasks in research institutions, universities, public agencies, and businesses. Survey participants were selected via snowball sampling starting with the internal staff at the KISTI and expanding through continuous recommendations from relevant agency personnel. A total of 134 responses was obtained, and after excluding 40 responses with incomplete or discontinued answers during the survey, the valid responses amounted to 94. To understand the respondents’ demographics, this study collected additional information on their affiliated organizations, highest education level, and years of work experience (Table 6).

Table 6. Summary of survey results and respondent status

Category	Number of Responses	Ratio (%)	Category	Number of Responses	Ratio (%)	Category	Number of Responses	Ratio (%)
Research institute	40	43%	bachelor	18	19%	1-5 years	21	22%
University	16	17%	master	40	43%	6-10 years	20	21%
Public institution	21	22%	doctor	33	35%	11-20 years	25	27%
Company	16	17%	etc.	3	3%	21 years or more	28	30%
Others	1	1%						
Sum	94	100%	Sum	94	100%	Sum	94	100%

About 65% of the respondents worked in public institutions and research organizations, with 78% holding a master's or doctoral degree and 57% having more than 11 years of work experience. Cronbach's alpha was used to verify the internal consistency and reliability of the survey results, showing values above 0.7 in all categories: technology, data, organization, strategy, and social influence. The lowest Cronbach's alpha coefficient (0.752) was reported in the technology category for T1-1 (R&D investment sector), while the highest coefficient (0.929) was in the data category for D2-2 (readiness sector). Moreover, removing specific items resulted in lower Cronbach's alpha coefficients, reinforcing internal consistency. Table 7 shows examples of Cronbach's alpha values in this study.

Table 7. Social influence category reliability verification (example)

Middle classification	Metrics	Cronbach's alpha when item is removed	Cronbach's alpha
11. Contribution to bridging the digital gap	I1-1. Verify that the organization's information services allow nondiscriminatory access to all users	-	.924
	I1-2. Verify that the organization's content allows nondiscriminatory access to all users	-	
12. Economic effect	I2-1. Level of awareness of the extent to which the institution's resources have contributed to the development of national science and technology	.894	.878
	I2-2. Level of awareness of the extent to which the organization's activities have contributed to the creation of patents, etc.	.810	
	I2-3. Level of awareness of the extent to which the organization's activities have contributed to national competitiveness, such as technology exports	.771	
13. Educational effect	I3-1. Measure the extent to which the organization's activities are perceived to have contributed to the provision of educational materials	-	.886
	I3-2. Measure the extent to which the organization's activities are perceived to have contributed to users' lifelong education	-	
14. Degree of data openness	I4-1. Measure whether an organization's activities are perceived as contributing to data openness	.866	.906
	I4-2. Measure the perceived level of data openness of an organization	.878	
	I4-3. Measure whether people perceive their organization to be sharing data well	.853	
15. Overall satisfaction	I5-1. Whether the evaluation reflects the user's overall level of satisfaction with the services provided by the institution	.729	.854
	I5-3. Level of awareness that user feedback, such as improving user inconveniences, is being properly implemented	.833	
	I5-4. Degree of positive perception of the organization's existence and services	.825	

4.1 Confirmatory factor analysis of the technology category

The technology category contained a total of 37 measurement indicators across 9 middle categories. AMOS 22 was used to construct a confirmatory factor analysis model for these (Figure 3).

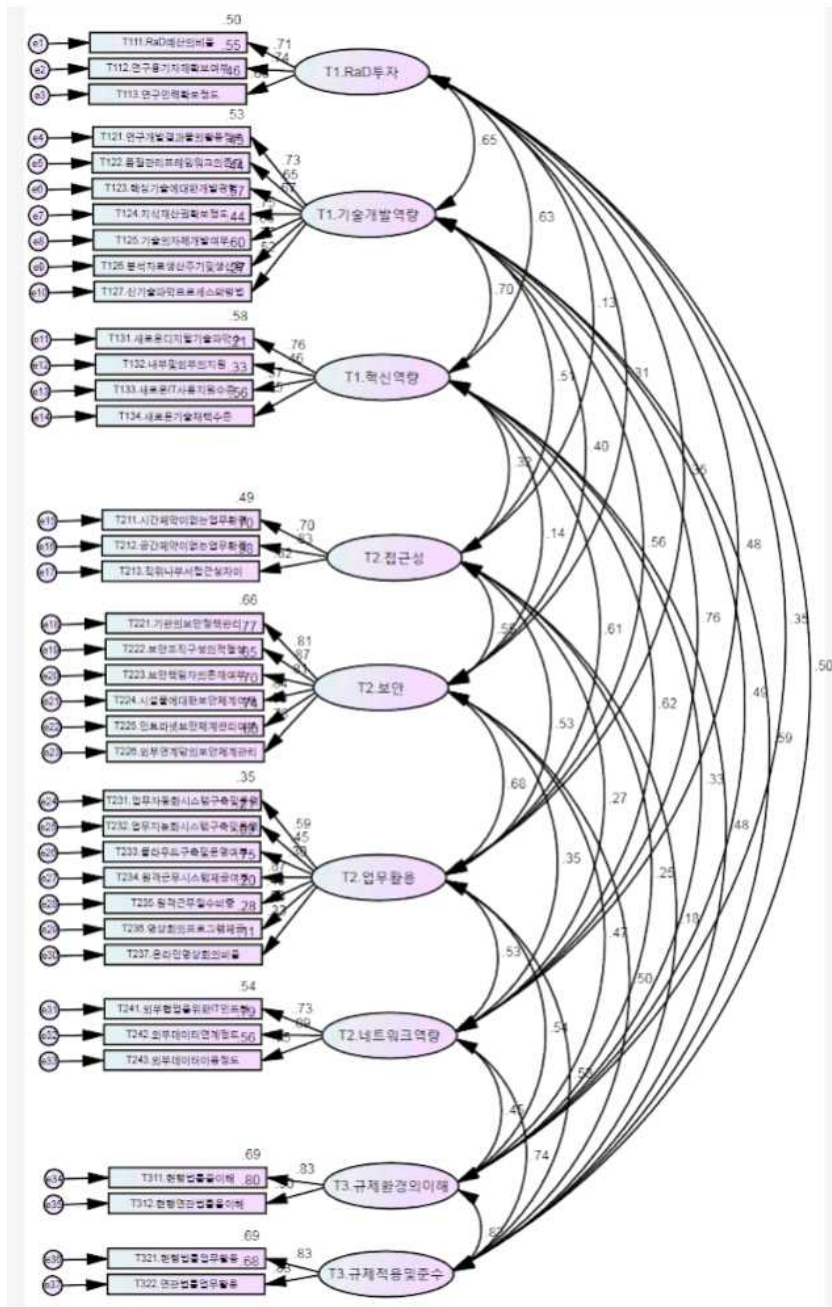


Fig. 3. Confirmatory factor analysis model of technical categories

During the first stage of validation, a comparison of unstandardized λ values, standard errors (SEs), and statistics including p-values within the technology category revealed that, based on the criterion of unstandardized λ values, all critical ratio (CR) values exceeded 1.96 ($p < 0.05$). Thus, all values met the criteria for the first stage.

Regarding content validity verification, the first step excluded measurement indicators with standardized λ values below 0.7. In addition, 15 measurement indicators were challenging to exclude because doing so would leave only one measurement indicator for accessibility and task utilization latent variables. This not only hindered the assessment of the relative importance of this singular indicator but also presented a statistical challenge in confirming discriminant validity. To address this, a further two measurement indicators within the accessibility and task utilization latent variables were excluded. Table 8 shows detailed information.

Table 8. Unstandardized and standardized coefficients of technical categories

Category		Unstandardized coefficients	SE.	CR.	P	Standardized coefficients
T1. R&D investment	---> T1-1-1. R&D budget ratio to total budget	0.982	0.179	5.492	***	0.709
T1. R&D investment	---> T1-1-2. Whether equipment necessary for research and development is secured	0.997	0.176	5.663	***	0.745
T1. R&D investment	---> T1-1-3. Degree of securing human resources required for research and development					0.68
T1. Technology development capability	---> T1-2-1. Degree of utilization of research and development results	1.072	0.174	6.161	***	0.726
T1. Technology development capability	---> T1-2-2. Existence of a quality management framework	0.899	0.16	5.626	***	0.654
T1. Technology development capability	---> T1-2-3. Existence of a development experience in core technologies					0.665
T1. Technology development capability	---> T1-2-4. Degree of securing rights to technology (intellectual property rights)	1.176	0.185	6.351	***	0.753
T1. Technology development capability	---> T1-2-5. Whether the technology introduced to work is developed independently	1.085	0.191	5.683	***	0.662
T1. Technology development capability	---> T1-2-6. Production cycle and production level of market information analysis data	1.216	0.188	6.482	***	0.772
T1. Technology development capability	---> T1-2-7. Existence of systematic processes and methodologies for identifying new technologies	0.686	0.15	4.561	***	0.518

Category		Unstandardized coefficients	SE.	CR.	P	Standardized coefficients
T1. Innovation capability	---> T1-3-1. Level of understanding of new digital technologies	1.412	0.272	5.2	***	0.763
T1. Innovation capability	---> T1-3-2. Evaluate efficient use of internal and external support	0.848	0.232	3.663	***	0.46
T1. Innovation capability	---> T1-3-3. Level of support for new IT 1 uses					0.573
T1. Innovation capability	---> T1-3-4. Level of adoption of new technologies	1.405	0.272	5.159	***	0.751
T2. Accessibility	---> T2-1-1. Existence of a work environment without time constraints	1.26	0.245	5.155	***	0.698
T2. Accessibility	---> T2-1-2. Existence of work environment without spatial restrictions	1.512	0.273	5.547	***	0.835
T2. Accessibility	---> T2-1-3. Existence of differences in 1 accessibility depending on position or department					0.617
T2. Security	---> T2-2-1. Existence of the institution's security policy (administrative, physical, intangible information and communication technology)	0.873	0.096	9.068	***	0.814
T2. Security	---> T2-2-2. Appropriateness of security organization composition	1.071	0.106	10.058	***	0.875
T2. Security	---> T2-2-3. Existence of a security officer 1					0.808
T2. Security	---> T2-2-4. Existence of a security system for facilities	0.959	0.101	9.473	***	0.839
T2. Security	---> T2-2-5. Internal Intranet security system management status and management level	0.93	0.095	9.804	***	0.86
T2. Security	---> T2-2-6. External link network security system management status and management level	0.901	0.106	8.481	***	0.776
T2. Work utilization	---> T2-3-1. Whether a work automation system is established and operated	1.877	0.715	2.625	0.009	0.588
T2. Work utilization	---> T2-3-2. Whether a work intelligence system is established and operated	1.747	0.719	2.43	0.015	0.453
T2. Work utilization	---> T2-3-3. Whether to build and operate 1 a cloud system					0.304
T2. Work utilization	---> T2-3-4. Whether a remote work system is provided	3.142	1.12	2.806	0.005	0.866
T2. Work utilization	---> T2-3-5. Proportion of remote working days to total working days	1.612	0.666	2.421	0.015	0.449
T2. Work utilization	---> T2-3-6. Availability of online conference program	2.04	0.798	2.556	0.011	0.531
T2. Work utilization	---> T2-3-7. Ratio of online meetings to offline meetings	1.156	0.547	2.111	0.035	0.327
T2. Network capability	---> T2-4-1. Whether to build IT infrastructure for collaboration with external organizations (or systems)	0.849	0.122	6.955	***	0.735
T2. Network capability	---> T2-4-2. Degree of linkage to external data (public data or external agency data)	1.221	0.147	8.29	***	0.89

Category		Unstandardized coefficients	SE.	CR.	P	Standardized coefficients
T2. Network capability	---> T2-4-3. Degree of usage of external data (public data or external agency data)	1				0.748
T3. Understanding the regulatory environment	---> T3-1-1. Measures your understanding of the current laws that govern the work	1				0.83
T3. Understanding the regulatory environment	---> T3-1-2. Measures understanding of current laws relevant to the work	1.2	0.131	9.172	***	0.896
T3. Regulatory application and compliance	---> T3-2-1. Measures whether current laws mainly targeting the relevant work are being utilized when carrying out work	1				0.832
T3. Regulatory application and compliance	---> T3-2-2. Measures whether laws related to the work are being utilized when carrying out work	1.091	0.12	9.12	***	0.827

*, **, *** in the statistical analysis table mean p < .05, p < .01, and p < .001, respectively.

Furthermore, AVE values and composite reliability (CR) values were examined, and the results showed that both were above 0.5 and 0.7, respectively, in all areas, confirming convergent validity. Table 9 provides detailed numerical values.

Table 9. Statistics and average variance extracted (AVE) and composite reliability (CR) values to verify the central validity of the technical scale

		Unstandardized coefficient	SE.	CR.	Standardized coefficient	AVE	Concept reliability
T1. R&D investment	---> T1-1-1. R&D budget ratio to total budget	1	-	-	0.695	0.65331449	0.78974659
	---> T1-1-2. Whether equipment necessary for research and development is secured	1.085	0.233	4.661	0.781		
T1. Technology development capability	---> T1-2-1. Degree of utilization of research and development results	1	-	-	0.757	0.56976555	0.79868323
	---> T1-2-4. Degree of securing rights to technology (intellectual property rights)	1.084	0.156	6.964	0.776		
	---> T1-2-6. Production cycle and production level of market information analysis data	1	0.156	6.408	0.709		
T1. Innovation capability	---> T1-3-1. Level of understanding of new digital technologies	1	-	-	0.806	0.65816432	0.79361902
	---> T1-3-4. Level of adoption of new technologies	0.937	0.162	5.793	0.748		

			Unstandardized coefficient	SE.	CR.	Standardized coefficient	AVE	Concept reliability
T2. Security	--->	T2-2-1. Existence of the institution's security policy (administrative, physical, intangible information and communication technology)	0.884	0.097	9.081	0.821	0.7168031	0.93814111
	--->	T2-2-2. Appropriateness of security organization composition	1.069	0.108	9.858	0.87		
	--->	T2-2-3. Existence of a security officer	1	-	-	0.805		
	--->	T2-2-4. Existence of a security system for facilities	0.955	0.103	9.258	0.832		
	--->	T2-2-5. Internal Intranet security system management status and management level	0.938	0.096	9.763	0.864		
	--->	T2-2-6. External link network security system management status and management level	0.913	0.107	8.51	0.783		
T2. Network capability	--->	T2-4-1. Whether to build IT infrastructure for collaboration with external organizations (or systems)	0.852	0.125	6.828	0.728	0.66735383	0.85624781
	--->	T2-4-2. Degree of linkage to external data (public data or external agency data)	1.258	0.153	8.218	0.904		
	--->	T2-4-3. Degree of linkage to external data (public data or external agency data)	1	-	-	0.738		
T3. Understanding the regulatory environment	--->	T3-1-1. Measures your understanding of the current laws that govern the work	1	-	-	0.891	0.80008329	0.88884018
	--->	T3-1-2. Measures understanding of current laws relevant to the work	1.043	0.115	9.053	0.836		
T3. Regulatory application and compliance	--->	T3-2-1. Measures whether current laws mainly targeting the relevant work are being utilized when carrying out work	1	-	-	0.856	0.74704141	0.85508596
	--->	T3-2-2. Measures whether laws related to the work are being utilized when carrying out work	1.031	0.113	9.087	0.804		

The third step confirmed discriminant validity through the correlation of measurement indicators among the conceptual constructs and whether the AVE values were greater than the square of the correlation coefficient. Table 10 summarizes the comparison between correlation coefficients (squared) among the conceptual constructs and AVE values. Table 11 presents the results of the (correlation coefficient $\pm 2*SE$) range.

Table 10. Comparison of correlation coefficient square and average variance extracted (AVE) values

	T1. R&D investment	T1. Technology development capacity	T1. Innovation Capability	T2. Security	T2. Network capability	T3. Understanding the regulatory environment	AVE	Concept reliability
T1. R&D investment	1.00						0.653	0.790
T1. Technology development capability	0.34	1.00					0.569	0.798
T1. Innovation capability	0.30	0.36	1.00				0.658	0.794
T2. Security	0.11	0.22	0.00	1.00			0.717	0.938
T2. Network capability	0.17	0.55	0.35	0.12	1.00		0.667	0.856
T3. Understanding the regulatory environment	0.14	0.19	0.12	0.17	0.20	1.00	0.800	0.889
T3. Regulatory application and compliance	0.25	0.31	0.21	0.24	0.53	0.72	0.747	0.855

Table 11. Comparison of results over a (correlation coefficient $\pm 2*$ standard error) range

		Correlation coefficient	SE.	2*SE.	Lower limit(-)	Upper limit(+)
T1. R&D investment	<--> T1. Technology development capability	0.585	0.071	0.142	0.443	0.727
T1. R&D investment	<--> T1. Innovation capability	0.548	0.07	0.140	0.408	0.688
T1. R&D investment	<--> T2. Security	0.336	0.064	0.128	0.208	0.464
T1. R&D investment	<--> T2. Network capability	0.418	0.063	0.126	0.292	0.544
T1. R&D investment	<--> T3. Understanding the regulatory environment	0.377	0.06	0.120	0.257	0.497
T1. R&D investment	<--> T3. Regulatory application and compliance	0.495	0.064	0.128	0.367	0.623
T1. Technology development capability	<--> T1. Innovation capability	0.598	0.082	0.164	0.434	0.762

		Correlation coefficient	SE.	2*SE.	Lower limit(-)	Upper limit(+)
T1. Technology development capability	<--> T2. Security	0.464	0.08	0.160	0.304	0.624
T1. Technology development capability	<--> T2. Network capability	0.741	0.089	0.178	0.563	0.919
T1. Technology development capability	<--> T3. Understanding the regulatory environment	0.441	0.071	0.142	0.299	0.583
T1. Technology development capability	<--> T3. Regulatory application and compliance	0.557	0.075	0.150	0.407	0.707
T1. Innovation capability	<--> T2. Security	0.064	0.071	0.142	-0.078	0.206
T1. Innovation capability	<--> T2. Network capability	0.594	0.081	0.162	0.432	0.756
T1. Innovation capability	<--> T3. Understanding the regulatory environment	0.341	0.07	0.140	0.201	0.481
T1. Innovation capability	<--> T3. Regulatory application and compliance	0.455	0.073	0.146	0.309	0.601
T2. Security	<--> T2. Network capability	0.349	0.074	0.148	0.201	0.497
T2. Security	<--> T3. Understanding the regulatory environment	0.408	0.074	0.148	0.260	0.556
T2. Security	<--> T3. Regulatory application and compliance	0.487	0.077	0.154	0.333	0.641
T2. Network capability	<--> T3. Understanding the regulatory environment	0.447	0.07	0.140	0.307	0.587
T2. Network capability	<--> T3. Regulatory application and compliance	0.728	0.082	0.164	0.564	0.892
T3. Understanding the regulatory environment	<--> T3. Regulatory application and compliance	0.848	0.08	0.160	0.688	1.008

The validation results showed that the square of the correlation coefficients between all conceptual structures was above the lower limit of the AVE values. However, the (correlation coefficient \pm 2*SE) range of T3. Understanding regulatory environment <--> T3. Regulatory application and compliance included 1. Measurement indicators related to regulations were extracted from the IMD World Competitiveness Center (2021) and Portulans Institute (2021) models. Although these indicators were initially divided into understanding and applying the regulatory environment for subcategories, discriminant validity was not confirmed. The term “regulatory environment” was integrated because it was also used in other studies. Table 12 presents the intermediate classifications and final measurement indicators for the technology category after confirmatory factor analysis.

Table 12. Final classification and measurement indicators for the technology category

Category	Metrics
T1. R&D investment	T1-1-1. R&D budget ratio to total budget
	T1-1-2. Whether equipment necessary for research and development is secured
T1. Technology development capability	T1-2-1. Degree of utilization of research and development results
	T1-2-4. Degree of securing rights to technology (intellectual property rights)
	T1-2-6. Production cycle and production level of market information analysis data
T1. Innovation capability	T1-3-1. Level of understanding of new digital technologies
	T1-3-4. Level of adoption of new technologies
T2. Security	T2-2-1. Existence of the institution's security policy (administrative, physical, intangible information and communication technology)
	T2-2-2. Appropriateness of security organization composition
	T2-2-3. Existence of a security officer
	T2-2-4. Existence of a security system for facilities
	T2-2-5. Internal Intranet security system management status and management level
	T2-2-6. External link network security system management status and management level
T2. Network capability	T2-4-1. Whether to build IT infrastructure for collaboration with external organizations (or systems)
	T2-4-2. Degree of linkage to external data (public data or external agency data)
	T2-4-3. Degree of linkage to external data (public data or external agency data)
T3. Regulatory environment	T3-1-1. Measures your understanding of the current laws that govern the work
	T3-1-2. Measures understanding of current laws relevant to the work
	T3-2-1. Measures whether current laws mainly targeting the relevant work are being utilized when carrying out work
	T3-2-2. Measures whether laws related to the work are being utilized when carrying out work

4.2 Confirmatory factor analysis of the data category

The data category had a total of 45 measurement indicators distributed across 10 subcategories. Table 13 shows the unstandardized λ values, SEs, p-values, and standardized λ values. After calculating the critical ratio (CR) based on the criterion of unstandardized λ values exceeding 1.96 ($p < 0.05$), and considering standardized λ values above 0.7, a total of 9 measurement indicators with standardized coefficients below 0.7 were excluded, as shaded in Table 13.

Table 13. Unstandardized and standardized coefficients of the data category

Category		Unstandardized coefficient	SE.	CR.	P	Standardized coefficient
D1. Diversity	----> D1-1-1. Degree of securing (collecting) data suitable for purpose	0.673	0.092	7.315	***	0.719
D1. Diversity	----> D1-1-2. Degree of securing (collecting) uniform and unbiased data	0.943	0.108	8.731	***	0.83
D1. Diversity	----> D1-1-3. Whether to remove biased data that may be included in the data	1	-	-		0.78
D1. Diversity	----> D1-1-4. Verify that actual environment and situation characteristics are reflected when acquiring data under an artificial environment	0.859	0.108	7.994	***	0.773
D1. Diversity	----> D1-1-5. Verify that the environment and conditions are consistent when acquiring data under an artificial environment	0.858	0.108	7.917	***	0.767
D1. Diversity	----> D1-1-6. (Provided) Whether various data are provided	0.57	0.127	4.477	***	0.465
D1. Compatibility	----> D1-2-1. (Accuracy) Measure the accuracy of logical model, identifier, physical structure, and attribute meaning	1.065	0.124	8.578	***	0.8
D1. Compatibility	----> D1-2-2. (Consistency) Measure whether data are consistently defined and agree with each other	1.046	0.117	8.967	***	0.829
D1. Compatibility	----> D1-2-3. (Validity) Measure whether a data item satisfies a defined validity range (e.g., does the format of the data meet the validity range or does the data meet the domain validity range?)	1	-	-		0.828
D1. Usability	----> D1-3-1. Reusable data and metadata	1	0.175	5.7	***	0.649
D1. Usability	----> D1-3-10. A model that satisfies user requirements	0.861	0.143	6.032	***	0.693
D1. Usability	----> D1-3-2. Recyclability data identification	1.022	0.161	6.33	***	0.734
D1. Usability	----> D1-3-3. CC license	1	-	-		0.651
D1. Usability	----> D1-3-4. Hardware or software environment	1.102	0.182	6.061	***	0.697
D1. Usability	----> D1-3-5. The ratio of the number of open format data to the total number of data	1.091	0.176	6.202	***	0.716
D1. Usability	----> D1-3-6. Original text and dataset acquisition and registration rate	1.238	0.194	6.383	***	0.742
D1. Usability	----> D1-3-7. Data and metadata	1.266	0.192	6.586	***	0.77
D1. Usability	----> D1-3-8. Essential information needed for the search	1.218	0.184	6.617	***	0.775
D1. Usability	----> D1-3-9. User requirements	1.226	0.169	7.27	***	0.872
D1. Timeliness	----> D1-4-1. Measure whether the appropriate update period is defined and implemented according to the nature of the data	1.033	0.119	8.686	***	0.789
D1. Timeliness	----> D1-4-2. Measure whether the data provided is up to date	0.937	0.111	8.407	***	0.77
D1. Timeliness	----> D1-4-3. Measure whether the acquired data is synchronized	1				0.814

Category		Unstandardized coefficient	SE.	CR.	P	Standardized coefficient
D1. Timeliness	----> D1-4-4. Measure whether the work time from receipt of information requirements to collection, processing, and provision is minimized	0.981	0.121	8.099	***	0.749
D1. Interoperability	----> D1-5-1. Measure whether data and metadata use a principled glossary	1.471	0.201	7.314	***	0.893
D1. Interoperability	----> D1-5-2. Measure compliance with standard domains and standard terminology	1.346	0.181	7.419	***	0.912
D1. Interoperability	----> D1-5-3. Measures whether a representative file format that is widely used is used	1	-	-		0.653
D1. Security	----> D1-6-1. Measure whether data ownership is in place	0.844	0.11	7.655	***	0.697
D1. Security	----> D1-6-2. Measure whether there is a data access restriction policy and it is being implemented	1.062	0.103	10.301	***	0.846
D1. Security	----> D1-6-3. Measure data protection level	1	-	-		0.861
D1. Maintainability	----> D1-7-1. Measure whether data has changed and history is managed	1.258	0.214	5.878	***	0.791
D1. Maintainability	----> D1-7-2. Measure the existence of a maintenance policy	1.445	0.235	6.145	***	0.853
D1. Maintainability	----> D1-7-3. Measure maintainability	1	-	-		0.604
D2. Readiness	----> D2-1-1. Measure whether laws and systems for data construction, management, and use, security, personal information protection, etc. have been sufficiently reviewed and reflected	1.037	0.098	10.552	***	0.832
D2. Readiness	----> D2-1-2. Measure whether there is a systematically established framework or process for data construction, management, and utilization	0.983	0.093	10.618	***	0.835
D2. Readiness	----> D2-1-3. Measure whether the organizational structure, roles, and responsibilities for data construction are systematically established and managed	1	-	-		0.867
D2. Readiness	----> D2-1-4. Measure whether plans are being established and managed for configuring tools and environments for data construction	1.044	0.091	11.5	***	0.871
D2. Readiness	----> D2-1-5. Measure whether quality monitoring processes and control procedures for data management are established and managed	1.071	0.098	10.935	***	0.849
D2. Completeness	----> D2-2-1. Measure whether the organization has a structure for building data that meets its initial purpose and goals	1	-	-		0.77
D2. Completeness	----> D2-2-2. Measure whether raw data collection methods, standards, training, and inspection are systematically established and managed	0.955	0.113	8.475	***	0.803

Category		Unstandardized coefficient	SE.	CR.	P	Standardized coefficient
D2. Completeness	----> D2-2-3. Measure whether the purification of collected raw data is systematically established, managed, and performed	0.949	0.099	9.55	***	0.883
D2. Completeness	----> D2-2-4. Measure whether data is stored for ease of distribution and use	0.737	0.106	6.969	***	0.683
D2. Usefulness	----> D2-3-1. Measure whether the organization's requirements are sufficiently reflected in the construction process	1	-	-		0.742
D2. Usefulness	----> D2-3-2. Measure whether the scope and detail of data is suitable for the purpose	1.028	0.135	7.597	***	0.763
D2. Usefulness	----> D2-3-3. Measure whether the results according to the construction model satisfy the target performance indicators	1.276	0.153	8.339	***	0.83
D2. Usefulness	----> D2-3-4. Measure whether there is an alternative in case the intended target or performance indicator is not met	1.242	0.173	7.186	***	0.726

In the statistical analysis table, *** means $p < .001$.

Category		Unstandardized coefficient	SE.	CR.	P	Standardized coefficient
D1. Diversity	----> D1-1-1. Degree of securing (collecting) data suitable for purpose	0.673	0.092	7.315	***	0.719
D1. Diversity	----> D1-1-2. Degree of securing (collecting) uniform and unbiased data	0.943	0.108	8.731	***	0.83
D1. Diversity	----> D1-1-3. Whether to remove biased data that may be included in the data	1	-	-		0.78
D1. Diversity	----> D1-1-4. Verify that actual environment and situation characteristics are reflected when acquiring data under an artificial environment	0.859	0.108	7.994	***	0.773
D1. Diversity	----> D1-1-5. Verify that the environment and conditions are consistent when acquiring data under an artificial environment	0.858	0.108	7.917	***	0.767
D1. Diversity	----> D1-1-6. (Provided) Whether various data are provided	0.57	0.127	4.477	***	0.465
D1. Compatibility	----> D1-2-1. (Accuracy) Measure the accuracy of logical model, identifier, physical structure, and attribute meaning	1.065	0.124	8.578	***	0.8
D1. Compatibility	----> D1-2-2. (Consistency) Measure whether data are consistently defined and agree with each other	1.046	0.117	8.967	***	0.829
D1. Compatibility	----> D1-2-3. (Validity) Measure whether a data item satisfies a defined validity range (e.g., does the format of the data meet the validity range or does the data meet the domain validity range?)	1	-	-		0.828
D1. Usability	----> D1-3-1. Reusable data and metadata	1	0.175	5.7	***	0.649
D1. Usability	----> D1-3-10. A model that satisfies user requirements	0.861	0.143	6.032	***	0.693

Category		Unstandardized coefficient	SE.	CR.	P	Standardized coefficient
D1. Usability	---> D1-3-2. Recyclability data identification	1.022	0.161	6.33	***	0.734
D1. Usability	---> D1-3-3. CC license	1	-	-		0.651
D1. Usability	---> D1-3-4. Hardware or software environment	1.102	0.182	6.061	***	0.697
D1. Usability	---> D1-3-5. The ratio of the number of open format data to the total number of data	1.091	0.176	6.202	***	0.716
D1. Usability	---> D1-3-6. Original text and dataset acquisition and registration rate	1.238	0.194	6.383	***	0.742
D1. Usability	---> D1-3-7. Data and metadata	1.266	0.192	6.586	***	0.77
D1. Usability	---> D1-3-8. Essential information needed for the search	1.218	0.184	6.617	***	0.775
D1. Usability	---> D1-3-9. User requirements	1.226	0.169	7.27	***	0.872
D1. Timeliness	---> D1-4-1. Measure whether the appropriate update period is defined and implemented according to the nature of the data	1.033	0.119	8.686	***	0.789
D1. Timeliness	---> D1-4-2. Measure whether the data provided is up to date	0.937	0.111	8.407	***	0.77
D1. Timeliness	---> D1-4-3. Measure whether the acquired data is synchronized	1				0.814
D1. Timeliness	---> D1-4-4. Measure whether the work time from receipt of information requirements to collection, processing, and provision is minimized	0.981	0.121	8.099	***	0.749
D1. Interoperability	---> D1-5-1. Measure whether data and metadata use a principled glossary	1.471	0.201	7.314	***	0.893
D1. Interoperability	---> D1-5-2. Measure compliance with standard domains and standard terminology	1.346	0.181	7.419	***	0.912
D1. Interoperability	---> D1-5-3. Measures whether a representative file format that is widely used is used	1	-	-		0.653
D1. Security	---> D1-6-1. Measure whether data ownership is in place	0.844	0.11	7.655	***	0.697
D1. Security	---> D1-6-2. Measure whether there is a data access restriction policy and it is being implemented	1.062	0.103	10.301	***	0.846
D1. Security	---> D1-6-3. Measure data protection level	1	-	-		0.861
D1. Maintainability	---> D1-7-1. Measure whether data has changed and history is managed	1.258	0.214	5.878	***	0.791
D1. Maintainability	---> D1-7-2. Measure the existence of a maintenance policy	1.445	0.235	6.145	***	0.853
D1. Maintainability	---> D1-7-3. Measure maintainability	1	-	-		0.604
T2. Network capability	---> T2-4-3. Degree of usage of external data (public data or external agency data)	1				0.748
T3. Understanding the regulatory environment	---> T3-1-1. Measures your understanding of the current laws that govern the work	1				0.83

Category			Unstandardized coefficient	SE.	CR.	P	Standardized coefficient
T3. Understanding the regulatory environment	--->	T3-1-2. Measures understanding of current laws relevant to the work	1.2	0.131	9.172	***	0.896
T3. Regulatory application and compliance	--->	T3-2-1. Measures whether current laws mainly targeting the relevant work are being utilized when carrying out work	1				0.832
T3. Regulatory application and compliance	--->	T3-2-2. Measures whether laws related to the work are being utilized when carrying out work	1.091	0.12	9.12	***	0.827

After checking the AVE values and conceptual reliability, concentration validity was confirmed in all areas with values of 0.5 and above and 0.7 and above, respectively. However, comparing the squares of the correlation coefficients with the AVE values, except for diversity \Leftrightarrow utility, utility \Leftrightarrow suitability, and all other structural concepts, timeliness \Leftrightarrow readiness, completeness, utility, interoperability \Leftrightarrow utility, security, maintenance, maintenance \Leftrightarrow readiness, completeness, security, completeness \Leftrightarrow utility, in a total of 18 areas, AVE values exceeded the upper limit, indicating inadequate discriminant validity. A reevaluation of the (correlation coefficient $\pm 2*SE$) ranges showed that discriminant validity was not confirmed in 9 areas, including utility \Leftrightarrow timeliness, interoperability, completeness, utility, timeliness \Leftrightarrow readiness, interoperability \Leftrightarrow maintenance, security, maintenance \Leftrightarrow security, completeness \Leftrightarrow security.

Even if verification were conducted within the (correlation coefficient $\pm 2*SE$) range, the extremely high correlation in at least nine areas means that the differences between the structural concepts could not be confirmed statistically. However, a review of these structural concepts shows that even though they are not the same, they may have a high correlation because they are related. For example, interoperability and utility showed a higher correlation coefficient than the threshold. While interoperability and utility are distinct, high interoperability naturally leads to high utility, explaining the high correlation coefficient. In the initial digital maturity assessment model developed in this study, the structural concepts were derived from concepts and metrics used in various fields and studies (utility: ISO/IEC 25012; RDA FAIR Data Maturity Model WG (2020); timeliness: ISO/IEC 9126; ISO/IEC 25012; interoperability: RDA FAIR Data Maturity Model WG (2020); completeness: NIA (2021, 2022); TTA.KO-10.1339; CMMI (Lanin, 2008); ISO 9001; usefulness: ISO 8000-61; NIA (2021, 2022); TTA.KO-10.1339; CMMI (Lanin, 2008); ISO 9001; timeliness: ISO/IEC 25012; ISO/IEC 9126; TTA.KO-10.1339; readiness: NIA (2018, 2021, 2022); ISO 8000-150; CMMI (Lanin, 2008); Kim, Lee & Lee (2017); maintainability: ISO/IEC 9126; security: NIA (2018); ISO/IEC 9126). Therefore, excluding or merging structural concepts based solely on high correlation might cause an evaluation model to fail to assess concepts commonly used in various models.

Furthermore, because these structural concepts are deemed essential in the data category, they were not excluded or merged in this model. Table 14 shows the final subcategories and metrics for the data category.

Table 14. Final classification and metrics of the data category

Category	Metrics
D1. Diversity	D1-1-1. Degree of securing (collecting) data suitable for purpose
	D1-1-2. Degree of securing (collecting) uniform and unbiased data
	D1-1-3. Whether to remove biased data that may be included in the data
	D1-1-4. Verify that actual environment and situation characteristics are reflected when acquiring data under an artificial environment
	D1-1-5. Verify that the environment and conditions are consistent when acquiring data under an artificial environment
D1. Compatibility	D1-2-1. (Accuracy) Measure the accuracy of logical model, identifier, physical structure, and attribute meaning
	D1-2-2. (Consistency) Measure whether data are consistently defined and agree with each other
	D1-2-3. (Validity) Measure whether a data item satisfies a defined validity range (e.g., does the format of the data meet the validity range or does the data meet the domain validity range?)
D1. Usability	D1-3-2. (Recyclability) Measure whether the data is highly identifiable
	D1-3-5. (Availability) Measure the ratio of the number of open format data to the total number of data
	D1-3-6. (Availability) Measure original text and dataset acquisition and registration rate
	D1-3-7. (Searchability) Measure whether unique and permanent identifiers are assigned to data and metadata
	D1-3-8. (Searchability) Measure whether essential information required for search (language diversity, topic information, etc.) is included
D1. Timeliness	D1-3-9. (Utility) Measure whether a data set is provided that can satisfy user requirements
	D1-4-1. Measure whether the appropriate update period is defined and implemented according to the nature of the data
	D1-4-2. Measure whether the data provided is up to date
	D1-4-3. Measure whether the acquired data is synchronized
	D1-4-4. Measure whether the work time from receipt of information requirements to collection, processing, and provision is minimized
D1. Interoperability	D1-5-1. Measure whether data and metadata use a principled glossary
	D1-5-2. Measure compliance with standard domains and standard terminology
D1. Security	D1-6-2. Measure whether there is a data access restriction policy and it is being implemented
	D1-6-3. Measure data protection level
D1. Maintainability	D1-7-1. Measure whether data has changed and history is managed
	D1-7-2. Measure the existence of a maintenance policy
D2. Readiness	D2-1-1. Measure whether laws and systems for data construction, management, and use, security, personal information protection, etc. have been sufficiently reviewed and reflected
	D2-1-2. Measure whether there is a systematically established framework or process for data construction, management, and utilization
	D2-1-3. Measure whether the organizational structure, roles, and responsibilities for data construction are systematically established and managed
	D2-1-4. Measure whether plans are being established and managed for configuring tools and environments for data construction
	D2-1-5. Measure whether quality monitoring processes and control procedures for data management are established and managed

Category	Metrics
D2. Completeness	D2-2-1. Measure whether the organization has a structure for building data that meets its initial purpose and goals
	D2-2-2. Measure whether raw data collection methods, standards, training, and inspection are systematically established and managed
	D2-2-3. Measure whether the purification of collected raw data is systematically established, managed, and performed
	D2-2-4. Measure whether data is stored for ease of distribution and use
D2. Usefulness	D2-3-1. Measure whether the organization's requirements are sufficiently reflected in the construction process
	D2-3-2. Measure whether the scope and detail of data is suitable for the purpose
	D2-3-3. Measure whether the results according to the construction model satisfy the target performance indicators
	D2-3-4. Measure whether there is an alternative in case the intended target or performance indicator is not met

4.3 Confirmatory factor analysis of the strategy category

The strategy category had a total of 18 measurement indicators distributed across 8 subcategories. Table 15 shows unstandardized λ values, SEs, p-values, and standardized λ values. After confirming that the unstandardized λ values met the critical ratio (CR) condition exceeding 1.96 ($p < 0.05$), and considering standardized λ values above 0.7, no measurement indicators in the strategy category were specifically excluded because all standardized λ values exceeded 0.7.

Table 15. Unstandardized and standardized coefficients of the strategy category

Category		Unstandardized coefficient	SE.	CR.	P	Standardized coefficient
S1. Vision and goals	-> S1-1-1. Measure whether the vision and goals for digital transformation across the organization have been established and formalized	0.946	0.097	9.782	***	0.83
S1. Vision and goals	-> S1-1-2. Measure whether members are aware of the vision and goals for digital transformation at the organizational level	1.071	0.094	11.447	***	0.917
S1. Vision and goals	-> S1-1-3. Measure the level of awareness of digital transformation among members within the organization	1	-	-		0.839
S1. Policy and process	-> S1-2-1. Measure whether they have policies in place that align with their organization-level digital transformation strategy	1.053	0.095	11.12	***	0.896
S1. Policy and process	-> S1-2-2. Measure whether they have and manage a system (process) that matches the organization-level digital transformation strategy	1	-	-		0.845

Category		Unstandardized coefficient	SE.	CR.	P	Standardized coefficient
S1. Process innovation	-> S1-3-1. Measure whether they have a strategy for automation and standardization at the organizational level	0.818	0.098	8.382	***	0.787
S1. Process innovation	-> S1-3-2. Measure whether the organization is aware of the importance of process innovation and has a strategy for it	1	-	-		0.824
S2. Policy and process	-> S2-1-1. Measure whether they have policies in line with their digital transformation strategy and whether work is managed accordingly	0.907	0.085	10.657	***	0.876
S2. Policy and process	-> S2-1-2. Measure whether they have a work system that matches their digital transformation strategy and whether work is progressing accordingly	0.887	0.085	10.404	***	0.863
S2. Policy and process	-> S2-1-3. Measure whether policies and processes are evaluated and improved	1	-	-		0.842
S2. Talent acquisition strategy	-> S2-2-1. Measure whether they have a talent acquisition strategy, policy and system in place	1.105	0.114	9.721	***	0.863
S2. Talent acquisition strategy	-> S2-2-2. Measure whether the importance of key personnel related to digital technology is recognized and reflected in strategy	1	-	-		0.845
S2. Commercialization strategy	-> S2-3-1. Measure whether internal standards such as strategic systems or guidelines for commercialization exist	1.189	0.135	8.78	***	0.92
S2. Commercialization strategy	-> S2-3-2. Measure whether guidelines for utilizing the results of the relevant R&D exist, if actual commercialization does not occur	1	-	-		0.751
S2. R&D strategy	-> S2-4-1. Measure whether internal regulations such as R&D strategy system or process exist	1.043	0.091	11.411	***	0.899
S2. R&D strategy	-> S2-4-2. Measure whether an improvement system such as evaluation or feedback on R&D exists	1	-	-		0.863
S2. Service strategy	-> S2-5-1. Measure whether a strategic system such as internal policies and guidelines related to user services exists	0.652	0.103	6.319	***	0.706
S2. Service strategy	-> S2-5-2. Measure whether a feedback system for users' services exists	1	-	-		0.953

In the statistical analysis table, *** means $p < .001$.

After verifying the AVE values and concept reliability, both of which exceeded 0.5 and 0.7, respectively, and ensured satisfactory convergent validity in all areas, a discriminant validity issue was identified when comparing the squares of correlation coefficients with AVE values. Specifically, in the R&D strategy \Leftrightarrow businessization strategy area, only the AVE value exceeded the upper limit, indicating inadequate discriminant validity. However, a reevaluation of the (correlation coefficient $\pm 2*SE$) range showed no issues. Table 16 shows the final middle categories and measurement indicators for the strategy category.

Table 16. Final classification and metrics of the strategy category

Category	
S1. Vision and goals	---> S1-1-1. Measure whether the vision and goals for digital transformation across the organization have been established and formalized
S1. Vision and goals	---> S1-1-2. Measure whether members are aware of the vision and goals for digital transformation at the organizational level
S1. Vision and goals	---> S1-1-3. Measure the level of awareness of digital transformation among members within the organization
S1. Policy and process	---> S1-2-1. Measure whether they have policies in place that align with their organization-level digital transformation strategy
S1. Policy and process	---> S1-2-2. Measure whether they have and manage a system (process) that matches the organization-level digital transformation strategy
S1. Process innovation	---> S1-3-1. Measure whether they have a strategy for automation and standardization at the organizational level
S1. Process innovation	---> S1-3-2. Measure whether the organization is aware of the importance of process innovation and has a strategy for it
S2. Policy and process	---> S2-1-1. Measure whether they have policies in line with their digital transformation strategy and whether work is managed accordingly
S2. Policy and process	---> S2-1-2. Measure whether they have a work system that matches their digital transformation strategy and whether work is progressing accordingly
S2. Policy and process	---> S2-1-3. Measure whether policies and processes are evaluated and improved
S2. Talent acquisition strategy	---> S2-2-1. Measure whether they have a talent acquisition strategy, policy, and system in place
S2. Talent acquisition strategy	---> S2-2-2. Measure whether the importance of key personnel related to digital technology is recognized and reflected in strategy
S2. Commercialization strategy	---> S2-3-1. Measure whether internal standards such as strategic systems or guidelines for commercialization exist
S2. Commercialization strategy	---> S2-3-2. Measure whether guidelines for utilizing the results of the relevant R&D exist, if actual commercialization does not occur
S2. R&D strategy	---> S2-4-1. Measure whether internal regulations such as R&D strategy system or process exist
S2. R&D strategy	---> S2-4-2. Measure whether an improvement system such as evaluation or feedback on R&D exists
S2. Service strategy	---> S2-5-1. Measure whether a strategic system such as internal policies and guidelines related to user services exists
S2. Service strategy	---> S2-5-2. Measure whether a feedback system for users' services exists

4.4 Confirmatory factor analysis of the organization category

The organization category had a total of 35 measurement indicators distributed across 12 subcategories. Table 17 presents the unstandardized λ values, SEs, p-values, and standardized λ values. After confirming that the unstandardized λ values met the critical ratio (CR) condition exceeding 1.96 ($p < 0.05$), and considering standardized λ values above 0.7, six measurement indicators in the strategy category were excluded for having standardized λ values below 0.7.

Table 17. Unstandardized and standardized coefficients of the organization category

Category		Unstandardized coefficient	SE.	CR.	P	Standardized coefficient
O1. Composition of dedicated organization	-> O1-1-1. Measure whether a dedicated research and development organization for digital transformation technology is established and operated	1.08	0.128	8.468	***	0.858
O1. Composition of dedicated organization	-> O1-1-2. Measure whether a management system 1 such as policies and guidelines exists, if a core organization is organized and operating					0.866
O1. Organizational personnel composition	-> O1-2-1. Measure whether staffing is done according to appropriate standards	0.948	0.115	8.234	***	0.802
O1. Organizational personnel composition	-> O1-2-2. Measure whether the number of people 1 and managers is adequate					0.862
O1. Organizational skills	-> O1-3-1. Measure whether a communication system exists for maximizing the organization's own technical capabilities and for feedback	1.105	0.117	9.482	***	0.863
O1. Organizational skills	-> O1-3-2. Measure whether the organization is 1 comprised of members willing to learn new technologies					0.797
O1. Organizational skills	-> O1-3-3. Measure whether members have the 1.2 ability to absorb external knowledge and technology		0.128	9.367	***	0.855
O1. Organizational linkage	-> O1-4-1. Measure whether an interdepartmental 1.065 collaboration system exists		0.149	7.151	***	0.807
O1. Organizational linkage	-> O1-4-2. Measure whether formalized 1 decision-making procedures exist					0.696
O1. Organizational linkage	-> O1-4-3. Measure whether a mediator exists for 0.996 communication and cooperation		0.16	6.24	***	0.697
O2. Work initiative	-> O2-1-1. Measure the level of awareness of 0.773 organizational members' work initiative		0.093	8.336	***	0.756
O2. Work initiative	-> O2-1-2. Measure whether they are clearly aware 1 of their current work and participate autonomously					0.837
O2. Work initiative	-> O2-1-3. Measure whether they are aware of the 1.064 difficulties and areas for improvement in their current work		0.102	10.432	***	0.881
O2. Work resilience	-> O2-3-1. Measure their willingness to accept 0.892 changes in your work		0.096	9.276	***	0.768
O2. Work resilience	-> O2-3-2. Measure whether a communication 1 channel exists to relieve work stress					0.831

Category		Unstandardized coefficient	SE.	CR.	P	Standardized coefficient
O2. Readiness for change	-> O2-4-1. Measure the readiness for new work changes	1.361	0.152	8.958	***	0.909
O2. Readiness for Change	-> O2-4-2. Measure the level of expectations for the new environment that comes with work change					0.725
O2. Technology management skills	-> O2-5-1. Measure whether individuals devote resources (energy) to adopt new technologies	0.961	0.113	8.511	***	0.762
O2. Technology management skills	-> O2-5-2. Measure how much influence management of individual capabilities is perceived to have on the organization	1				0.83
O3. Leadership system	-> O3-1-1. Measure whether there is a chief data officer or chief analytics officer (CAO)	0.991	0.134	7.394	***	0.67
O3. Leadership system	-> O3-1-2. Measure whether formal procedures for decision-making exist					0.876
O3. Leadership system	-> O3-1-3. Measure whether decisions are perceived as being made appropriately	0.992	0.106	9.364	***	0.783
O3. Executive CIO Role	-> O3-2-1. Measure the level of awareness of members regarding the work capabilities of management (CIO)	1.159	0.172	6.751	***	0.757
O3. Executive CIO role	-> O3-2-2. Measure how much management recognizes the importance of digital strategy	1				0.68
O3. Executive CIO role	-> O3-2-3. Measure whether a communication system (superior/subordinate) exists centered on management	1.338	0.189	7.063	***	0.796
O3. Executive CIO role	-> O3-2-4. Measure whether an organization-level system exists to develop management capabilities	1.393	0.197	7.073	***	0.797
O4. Workforce management	-> O4-1-1. Measure the overall level of awareness of the human resources management system	0.961	0.096	10.046	***	0.804
O4. Workforce management	-> O4-1-2. Measure whether there is an internal evaluation system at the organizational level for the work capabilities of each individual/management of the organization					0.88
O4. Workforce management	-> O4-1-3. Measure whether a sufficient communication system exists related to human resource management	0.944	0.092	10.286	***	0.815
O4. Workforce management	-> O4-1-4. Measure the level of awareness of the fairness of the human resource management system	0.917	0.102	9.015	***	0.755
O4. Workforce management	-> O4-1-5. Turnover rate/exit rate (%)	0.867	0.156	5.57	***	0.535
O4. Talent education	-> O4-2-1. Measure the types of education and training (number of cases/year)	1.026	0.056	18.255	***	0.955
O4. Talent education	-> O4-2-2. Measure the education and training development (number of cases/year)	1				0.949
O4. Talent education	-> O4-2-3. Measure the overall level of satisfaction with education and training	0.619	0.096	6.468	***	0.578
O4. Talent education	-> O4-2-4. Measure whether a channel exists for feedback and communication about education and training	0.787	0.089	8.883	***	0.708

In the statistical analysis table, *** means $p < .001$.

After confirming the AVE values and concept reliability, which were above 0.5 and 0.7, respectively, indicating satisfactory convergent validity in all areas, issues in discriminant validity were identified. Using the criterion of discriminant validity verification within the (correlation coefficient $\pm 2*SE$) range, problems were reported in seven areas: work resilience \Leftrightarrow organizational technical competence, change preparedness, personnel management, task leadership, change preparedness \Leftrightarrow technical management competence, leadership role of management CIO \Leftrightarrow leadership system, talent management.

Despite the high correlation coefficients in these areas, they represent different concepts that have been used in various studies (Heo & Cheon, 2021; Hong, Choi & Kim, 2019; Gartner, n.d.; IMD World Competitiveness Center, 2021; IMPULS, n.d.; Portulans Institute, 2021; Singapore Economic Development Board, 2020). Considering potential issues in the completeness of the evaluation model, this study did not exclude or integrate these areas. Table 18 shows the final subcategories and measurement indicators for the organization category.

Table 18. Final classification and metrics of the organization category

Category	Metrics
O1. Composition of dedicated organization	O1-1-1. Measure whether a dedicated research and development organization for digital transformation technology is established and operated
	O1-1-2. Measure whether a management system such as policies and guidelines exists, if a core organization is organized and operating
O1. Organizational personnel composition	O1-2-1. Measure whether staffing is done according to appropriate standards
	O1-2-2. Measure whether the number of people and managers is adequate
O1. Organizational skills	O1-3-1. Measure whether a communication system exists for maximizing the organization's own technical capabilities and for feedback
	O1-3-2. Measure whether the organization is comprised of members willing to learn new technologies
	O1-3-3. Measure whether members have the ability to absorb external knowledge and technology
O2. Work initiative	O2-1-1. Measure the level of awareness of organizational members' work initiative
	O2-1-2. Measure whether they are clearly aware of their current work and participate autonomously
	O2-1-3. Measure whether they are aware of the difficulties and areas for improvement in their current work
O2. Work resilience	O2-3-1. Measure their willingness to accept changes in your work
	O2-3-2. Measure whether a communication channel exists to relieve work stress
O2. Readiness for change	O2-4-1. Measure the readiness for new work changes
	O2-4-2. Measure the level of expectations for the new environment that comes with work change
O2. Technology management skills	O2-5-1. Measure whether individuals devote resources (energy) to adopt new technologies
	O2-5-2. Measure how much influence management of individual capabilities is perceived to have on the organization
O3. Leadership system	O3-1-2. Measure whether formal procedures for decision-making exist
	O3-1-3. Measure whether decisions are perceived as being made appropriately

Category	Metrics
O3. Executive CIO Role	O3-2-1. Measure the level of awareness of members regarding the work capabilities of management (CIO)
	O3-2-3. Measure whether a communication system (superior/subordinate) exists centered on management
	O3-2-4. Measure whether an organization-level system exists to develop management capabilities
O4. Workforce management	O4-1-1. Measure the overall level of awareness of the human resources management system
	O4-1-2. Measure whether there is an internal evaluation system at the organizational level for the work capabilities of each individual/management of the organization
	O4-1-3. Measure whether a sufficient communication system exists related to human resource management
	O4-1-4. Measure the level of awareness of the fairness of the human resource management system
O4. Talent education	O4-2-1. Measure the types of education and training (number of cases/year)
	O4-2-2. Measure the education and training development (number of cases/year)
	O4-2-4. Measure whether a channel exists for feedback and communication about education and training

4.5 Confirmatory factor analysis of the social influence category

The social influence category had a total of 18 measurement indicators across 5 subcategories. Table 19 shows the unstandardized λ values, SEs, p-values, and standardized λ values. Following the criterion of unstandardized λ values with critical ratios (CR) exceeding 1.96 ($p < .05$) and considering standardized λ values above 0.7, one measurement indicator with a standardized λ value below 0.7 was identified and subsequently excluded from the social influence category.

Table 19. Unstandardized and standardized coefficients of the social influence category

Category	Unstandardized coefficient	SE.	CR.	P	Standardized coefficient	
I-1. Contribute to bridging the digital gap	-> I-1-1. Verify that the organization's information services allow non-discriminatory access to all users	0.96	0.074	12.932	***	0.912
I-1. Contribute to bridging the digital gap	-> I-1-2. Verify that the organization's contents allow non-discriminatory access to all users		-	-		0.941
I-2. Economic effect	-> I-2-1. Level of awareness of the extent to which the institution's resources have contributed to national scientific and technological development	0.902	0.105	8.623	***	0.76
I-2. Economic effect	-> I-2-2. Level of awareness of the extent to which the organization's activities have contributed to the creation of patents, etc.		-	-		0.857

Category		Unstandardized coefficient	SE.	CR.	P	Standardized coefficient
I-2. Economic effect	-> I-2-3. Level of awareness of the extent to which the organization's activities have contributed to national competitiveness, including technology exports, etc.	1.109	0.1	11.126	***	0.922
I-3. Educational effect	-> I-3-1. Measures the extent to which the organization's activities are perceived to have contributed to the provision of educational materials	1.135	0.143	7.958	***	0.942
I-3. Educational effect	-> I-3-2. Measures the extent to which the organization's activities are perceived to have contributed to users' lifelong education	1	-	-		0.845
I-4. Degree of data openness	-> I-4-1. Measures whether an organization's activities are perceived as contributing to data openness	0.949	0.088	10.845	***	0.88
I-4. Degree of data openness	-> I-4-2. Measures the perceived level of data openness of an organization	1	-	-		0.846
I-4. Degree of data openness	-> I-4-3. Measures whether organizations perceive data sharing to be good	0.974	0.087	11.153	***	0.896
I-5. Overall satisfaction	-> I-5-1. Whether the evaluation reflects the user's overall level of satisfaction with the services provided by the institution	2.63	1.281	2.054	0.04	0.877
I-5. Overall satisfaction	-> I-5-2. Whether the assessment reflects the user's overall level of satisfaction with the institution's system	1	-	-		0.221
I-5. Overall satisfaction	-> I-5-3. Whether it is recognized that user feedback, such as resolving user inconveniences, is being properly provided	2.431	1.197	2.03	0.042	0.747
I-5. Overall satisfaction	-> I-5-4. Degree of positive perception of the organization's existence and services	2.799	1.367	2.048	0.041	0.836

In the statistical analysis table, *** means $p < .001$.

After confirming the AVE values and concept reliability, all areas were confirmed to have achieved concentration validity, with values exceeding 0.5 and 0.7, respectively. Additionally, a comparison of the squares of the correlation coefficients with the AVE values showed that the AVE values in all areas were below the lower limit, indicating appropriate discriminant validity. Table 20 outlines the final middle categories and measurement indicators for the social influence category.

Table 20. Final classification and measurement indicators of the social influence category

Category	Metrics
I-1. Contribute to bridging the digital gap	I-1-1. Verify that the organization's information services allow non-discriminatory access to all users
	I-1-2. Verify that the organization's contents allow non-discriminatory access to all users

Category	Metrics
I-2. Economic effect	I-2-1. Level of awareness of the extent to which the institution’s resources have contributed to national scientific and technological development
	I-2-2. Level of awareness of the extent to which the organization’s activities have contributed to the creation of patents, etc.
	I-2-3. Level of awareness of the extent to which the organization’s activities have contributed to national competitiveness, including technology exports, etc.
I-3. Educational effect	I-3-1. Measures the extent to which the organization’s activities are perceived to have contributed to the provision of educational materials
	I-3-2. Measures the extent to which the organization’s activities are perceived to have contributed to users’ lifelong education
I-4. Degree of data openness	I-4-1. Measures whether an organization’s activities are perceived as contributing to data openness
	I-4-2. Measures the perceived level of data openness of an organization
	I-4-3. Measures whether organizations perceive data sharing to be good
I-5. Overall satisfaction	I-5-1. Whether the evaluation reflects the user’s overall level of satisfaction with the services provided by the institution
	I-5-3. Whether it is recognized that user feedback, such as resolving user inconveniences, is being properly provided
	I-5-4. Degree of positive perception of the organization’s existence and services
	I-5-4. Degree of positive perception of the organization’s existence and services

Preliminary model

Measure			Indicators
Major categories (5)	Medium category (16)	Subcategories (40)	Evaluation indicators corresponding to subcategories
<ul style="list-style-type: none"> • Technology • Data • Strategy • Organization(personnel) • (social)Impact 	<ul style="list-style-type: none"> • Technology(3) • Data(2) • Strategy(2) • Organization(personnel)(4) • (social)Impact (5) 	<ul style="list-style-type: none"> • Technology (9) • Data(10) • Strategy(8) • Organization(personnel)(13) • (social)Impact (0) 	<ul style="list-style-type: none"> • Technology (37) • Data (45) • Strategy (18) • Organization(personnel) (36) • (social)Impact (14)

Final model

Measure			Indicators
Major categories (5)	Medium category (16)	Subcategories (40)	Evaluation indicators corresponding to subcategories
<ul style="list-style-type: none"> • Technology • Data • Strategy • Organization(personnel) • (social)Impact 	<ul style="list-style-type: none"> • Technology(3) • Data(2) • Strategy(2) • Organization(personnel)(4) • (social)Impact (5) 	<ul style="list-style-type: none"> • Technology (9) • Data(10) • Strategy(8) • Organization(personnel)(13) • (social)Impact (0) 	<ul style="list-style-type: none"> • Technology (20, 17 indicators eliminated) • Data(36, 9 indicators eliminated) • Strategy(18, maintain) • Organization(personnel) (30, 6 indicators eliminated) • (social)Impact(13, 1 indicator eliminated)

Fig. 4. shows a comparison between the final and preliminary models

5. AHP analysis for weighting digital curation maturity model indicators

This study employed the AHP technique to derive relative weights for elements within the main and middle categories of the model, aiming for its objective use. Specific steps in AHP, such as the detailed scoring and feedback phases, were excluded as they were deemed inappropriate; instead, the focus was on general procedures, particularly weighting and consistency testing.

5.1 Survey questionnaire design

To measure weights within the main and middle categories of the digital maturity assessment model in line with the research objectives, a questionnaire was developed to allow individual evaluators to perform pairwise comparisons that indicate the relative importance of or preference between classification items. A total of 31 questions were presented, consisting of 10 questions for the 5 categories within the major classification and 21 questions for the 16 categories within the subclassification. The survey questions pertained to the relative importance of the major and sub-categories outlined in the digital transformation maturity model as illustrated in Figure 4. Respondents were instructed to rate the relative importance of two items for each question as “very important,” “important,” “similar,” “not important,” or “not very important.”

The survey was conducted using the SurveyMonkey platform, a professional survey service, from September 22 to 26, 2022. Its respondents were the participants in the confirmatory factor analysis survey. A total of 48 individuals responded. Tables 21 to 23 provide details regarding the demographic distribution of respondents according to affiliation type, highest educational attainment, and years of work experience.

Table 21. Respondents by organization type

Category	Number of responses	Ratio (%)
Research institute	21	46%
University	9	13%
Public institution	14	30%
Company	4	11%
Sum	48	100%

Among the respondents, 46% were affiliated with research institutions, representing an absolute majority. Including users who belonged to public institutions, this figure increased to 76%. In terms of the highest educational attainment, master’s degrees accounted for the largest proportion at 60%, and when doctoral degrees were included, it amounted to 92%. Additionally, although no significant variations were observed in the respondents’ work experience, the range of 11-20 years constituted the largest proportion at 49%. Overall, 69% of the respondents had more than 11 years of work experience.

Table 22. Respondents by educational attainment

Category	Number of responses	Ratio
Bachelor's	7	8%
Master's	27	60%
Doctoral	14	32%
Sum	48	100%

Table 23. Respondents by years of service

Category	Number of responses	Ratio
1-5 yrs	8	13%
6-10 yrs	11	18%
11-20 yrs	19	49%
21 yrs or more	10	20%
Sum	48	100%

5.2 Weight measurement

To measure the weights, the opinions of all respondents regarding each response must be consolidated. To this end, this study applied the geometric mean method, which calculates the geometric mean of responses from participants satisfying all consistency indices for each response item. The resulting geometric mean is then considered the overall opinion of all respondents, making it a frequently used method when universally assuming respondents' expertise (Yoo, 2012).

Table 24. Symmetric matrix of weight ratios

	Technology (T)	Data (D)	Strategy (S)	Organization (O)	Social impact (I)
Technology (T)	1	1.522222222	1.511111111	1.455555556	1.462962963
Data (D)	0.656934307	1	1.685185185	1.666666667	1.622222222
Strategy (S)	0.661764706	0.593406593	1	1.488888889	1.411111111
Organization (O)	0.687022901	0.6	0.671641791	1	1.333333333
Social impact (I)	0.683544304	0.616438356	0.708661417	0.75	1

5.3 Consistency verification

A consistency check was conducted after the weights were determined. Despite the simplicity of AHP in performing pairwise comparisons, a consistency check is essential for precise results. Hence, this study used a nonstandard symmetric matrix and the multiplication of matrices between the weights and the average of those results to derive λ . Consistency was subsequently verified

through the consistency index (CI) and the consistency ratio (CR). If the CR value is below 0.1, the responses to matrix A are considered logically consistent (Choi, 2020). In this study, the λ value was 5.07254333, the CI value was 0.01813583, and the CR value was 0.01619271, with $n = 5$ (a situation with a matrix of size 5). The random index (RI) value (an average CI value for 100 randomly generated symmetric matrices for a given size n) was 1.12. Since the CR value was below 0.1, consistency was confirmed. Table 25 shows the detailed numerical values.

Table 25. Summary of weighted sums and consistency measures

	Technology (T)	Data (D)	Strategy (S)	Organization (O)	Social impact (I)	Weight	Weighted sum	Consistency measure
Technology (T)	1	1.522222222	1.511111111	1.455555556	1.462962963	0.267288823	1.361401431	5.09337209
Data (D)	0.656934307	1	1.685185185	1.666666667	1.622222222	0.242125517	1.23243654	5.09007293
Strategy (S)	0.661764706	0.593406593	1	1.488888889	1.411111111	0.187270729	0.948250042	5.06352514
Organization (O)	0.687022901	0.6	0.671641791	1	1.333333333	0.159519284	0.805934514	5.05227012
Social impact (I)	0.683544304	0.616438356	0.708661417	0.75	1	0.143795647	0.728105858	5.06347634

In the same way, weights were derived for the middle categories within each main category.

5.4 Result of deriving weights for the main and middle categories

Above, the measured and validated weights were organized based on major and subcategories, and final weights were derived by multiplying the weights of the major categories by those of the middle categories.

A comprehensive weight synthesis among major categories showed that the technology category was the most dominant, and within middle categories, research and development within the technology category was the most dominant. However, the final weight derivation results showed that the middle category with the highest weight was data quality in the data category with a weight exceeding 0.16, followed by research and development in the technology category (0.12591379) and organizational strategy level in the strategy category (0.11704421) as dominant subcategory elements.

Table 26. Weighting results through the analytic hierarchy process of the digital maturity assessment model

Main category	Weight	Middle category	Weight	Final weight
Technology	0.26728882	R&D	0.47107763	0.12591379
		IT infrastructure	0.309989034	0.0828566
		Understanding and applying the regulatory environment	0.218933336	0.05851843
Data	0.24212552	Data quality	0.671532847	0.16259524
		Data management process	0.328467153	0.07953028
Strategy	0.18727073	Organizational strategy level	0.625	0.11704421
		Segmental strategy	0.375	0.07022652

Main category	Weight	Middle category	Weight	Final weight
(Social) influence	0.14379565	Organization	0.351705613	0.05610383
		Personal competency	0.232895061	0.03715125
		Leadership competency	0.232551713	0.03709648
		Operation and strategy	0.182847614	0.02916772
		Contribution to reducing the digital gap	0.251233934	0.03612635
		Economic effect	0.230732892	0.03317839
		Educational effect	0.180100732	0.0258977
		Data openness	0.185328748	0.02664947
		Satisfaction	0.152603694	0.02194375

6. Discussion

Although the validation process involved the participation of multiple experts, it does not guarantee the suitability of the results as a model for measuring digital transformation for several reasons. First, the professional criteria that the experts used to derive indicators and weights were difficult to standardize and had variable characteristics. Hence, further discussions must be conducted to determine whether factors such as workplace, tenure, field of work, and education can be considered clear criteria for expertise.

Second, the results validated through statistical devices might not adopt certain elements as indicators; nevertheless, these elements may still be considered valid in specific situations. Different results may be possible depending on the expert.

Lastly, although this study adopted AHP for weight assignment through pairwise comparisons, this approach is characterized by its ability to generate a large number of pairwise comparison items as the number of comparison subjects increases. Therefore, this study derived weights up to a mid-level classification. To enhance the model's practical utility, weight derivation must be extended to detailed items. Considering this, it might be worthwhile to explore options such as reducing the number of items, which may be performed by incorporating a high-order equation technique based on confirmatory factor analysis data rather than a direct application of AHP.

7. Conclusion and recommendation

The most fundamental function of the digital transformation maturity assessment model is to evaluate the current state of an organization, providing a means for control and suggesting actions for the future. It can serve as a framework for enhancing awareness and improvement in the analytical aspect, ensuring quality and reducing errors in the organization's key resources and services.

The model's significance rests in its reflection of different data types and data management processes and adoption of a digital transformation perspective. In addition, its evaluation criteria were structured

to allow for both quantitative and qualitative assessments. To help organizations establish evaluation criteria, the model categorizes them into technology, data, strategy, organization, and influence, making it possible to utilize them by sector. Importance is assigned to each criterion, which guides organizations in structuring evaluation items. The model is also meaningful in that it investigates several cases to derive its construction indicators. The derivation of weights for practical application also adds significance to the model.

After a careful consideration of the weight results, when evaluating the maturity of a digital curation institution from a digital transformation perspective, technological advancement is the most crucial factor, closely followed by data. Within the technological domain, this study emphasizes that the strength of research and development is more critical than the current state of IT infrastructure, reflecting expert insights into what is inherently crucial for technological advancement and highlighting the importance of technical investments.

In addition, data quality management significantly influences maturity assessment, highlighting the importance of not only having data but also ensuring well-managed and usable data.

Moreover, an organization's overall strategic responsiveness has a significant impact on maturity. Within the organizational domain, an organization's composition carries more weight than individual and leadership capabilities. This suggests that, for an optimal utilization of individual capabilities in digital curation maturity, organizational attention and structuring must focus on how the organization is configured.

With regard to social influence, this study revealed a prioritization of fundamental aspects such as addressing digital disparities and economic effects over systemic satisfaction, underscoring the significance of excelling in essential digital curation functions rather than merely meeting surface-level conditions for a positive impact on maturity assessment.

Digital transformation goals are challenging and cannot be easily attained within a short period. Moreover, the factors necessary for accomplishment are diverse and subject to change based on technological advancements and societal factors. The AHP analysis results in this paper shed light on what must be considered crucial when focusing on digital curation in institutions that manage and service digital knowledge information resources, especially in the digital transformation context.

Although AHP results are not absolute standards, this study indicates that technology and data are currently more critical to digital transformation than nontechnical factors, reflecting the perspectives of institutions that prepare and implement digital transformation. Despite the importance of balanced investments and execution across all factors, the results show that addressing challenges in technical aspects should take precedence.

In addition, while digital transformation is a crucial global challenge, identifying prominent success stories remains difficult. The models and metrics derived from this study must be tested in the field to determine their practical applicability for direct measurement. Furthermore, future studies must develop and refine new indicators tailored to each institution's objectives. Finally, the weights for each area and question must be investigated for the practical and rational use of this maturity model in the future.

Acknowledgements

This research was supported by the 2022 Basic Project of the Korea Institute of Science and Technology Information (KISTI) titled “Establishment of an Intelligent Science and Technology Information Curation System” (Project Number: K-22-L01-C01-S01).

References

- Choi, M. C. (2020). Evaluation of analytic hierarchy process method and development of a weight modified model. *Management & Information Systems Review*, 39(2), 145-162.
<https://doi.org/10.29214/damis.2020.39.2.009>
- Data Observation Network for Earth (n.d.). *DataONE best practices primer for data package creators*.
https://repository.oceanbestpractices.org/bitstream/handle/11329/502/DataONE_BP_Primer_02012.pdf
- Gartner. (n.d.). *Digital Government Maturity*. Retrieved April 21, 2025, from
<https://surveys.gartner.com/s/DigitalGovernmentMaturity>
- Heo, M., & Cheon, M. (2021). A Study on the Digital Transformation Readiness Through Developing and Applying Digital Maturity Diagnosis Model: Focused on the Case of a S Company in Oil and Chemical Industry. *Korean Management Review*, 50(1), 81-114.
<https://doi.org/10.17287/kmr.2021.50.1.81>
- Hong, S., Choi, Y., & Kim, G. (2019). A Study of Development of Digital Transformation Capacity. *Journal of The Korea Society of Information Technology Policy & Management*, 11(5), 1371-1381.
- IMD World Competitiveness Center. (2021). *IMD World Digital Competitiveness Ranking 2021*. Retrieved April 22, 2025, from
<https://investchile.gob.cl/wp-content/uploads/2022/03/imd-world-digital-competitiveness-rankings-2021.pdf>
- IMPULS. (n.d.). IMPULS Industrie 4.0-Readiness Online-Selbst-Check für Unternehmen. Retrieved April 21, 2025, from <https://www.iwconsult.de/projekte/industrie-40-readiness/>
- International Organization for Standardization & International Electrotechnical Commission. (2001). *Software engineering — Product quality — Part 1: Quality model (ISO/IEC 9126-1:2001)*.
- International Organization for Standardization & International Electrotechnical Commission. (2008). *Software engineering — Software product Quality Requirements and Evaluation (SQuARE) — Data quality model (ISO/IEC 25012:2008)*.
- International Organization for Standardization. (2016). *Data quality — Part 61: Data quality management: Process reference model (ISO 8000-61:2016)*.
- International Organization for Standardization. (2022). *Data quality — Part 150: Data quality management: Roles and responsibilities. (ISO 8000-150:2022)*.
- Jung, H. (2007). A Study of the Data Quality Evaluation. *Journal of Internet Computing and Services*, 8(4), 119-128.
-

- Kim, M., & Kim, M. (2020). Data quality control for data dams. *TTA Journal*, 192, 34-40.
- Korea Education and Research Information Service. (2021). *A plan to establish inclusive future education governance in response to digital transformation*. Daegu Metropolitan City: Korea Education and Research Information Service.
- Korea Institute of Public Administration. (2021). *Development and Utilization of Digital Level Diagnosis Model in Public Sector*. Seoul: Korea Institute of Public Administration.
- Lanin, I. (2008). Capability Maturity Model Integration (CMMI). Retrieved April 21, 2025, from <https://pt.slideshare.net/ivanlanin/capability-maturity-model-integrity-cmmi/6>
- National Information Society Agency. (2018). Public Data Quality Management Manual v2.0. Daegu Metropolitan City: Korea Information Society Agency.
- National Information Society Agency. (2021). Data Quality Management Guidelines for AI Training v1.0. Retrieved April 21, 2025, from <https://aihub.or.kr/aihubnews/qlityguidance/view.do?pageIndex=1&nttSn=10041&currMenu=&topMenu=&searchCondition=&searchKeyword=>
- National Information Society Agency. (2022). Data Quality Management Guidelines for AI Training v2.0. Retrieved April 21, 2025, from <https://aihub.or.kr/aihubnews/qlityguidance/view.do?currMenu=131&topMenu=103&nttSn=9831>
- National Library of Korea. (2021, September 28). A national library leading the digital transformation. <https://www.nl.go.kr/NL/contents/N50603000000.do?schM=view&id=40107&schBcid=normal0302>
- National Science and Technology Data Center Content Curation Center. (2020). *Establishment of science and technology content curation system*. Daejeon: Korea Institute of Science and Technology Information. Retrieved April 21, 2025, from <https://repository.kisti.re.kr/handle/10580/17347>
- Noh, K. (2019). *The proper methods of statistical analysis for dissertation*. Seoul: Hanbit Academy.
- OECD. (2019). Measuring the Digital Transformation: A ROADMAP FOR THE FUTURE. Retrieved April 22, 2025, from https://www.oecd.org/en/publications/measuring-the-digital-transformation_9789264311992-en.html
- Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64(1), 12-40.
- Park, S., & Cho, K. (2021). *The successful start of digital transformation*. Samsung SDS Insight Report. Retrieved April 21, 2025, from <https://www.samsungsds.com/kr/insights/dta.html>
- Portulans Institute. (2021, December 2). Network Readiness Index 2021. Retrieved April 21, 2025, from <https://networkreadinessindex.org/nri-2021-edition-press-release/>
- Principe, P., Manghi, P., Bardi, A., Vieira, A., Schirrwagen, J., & Pierrakos, D. (2019). A user journey in OpenAIRE services through the lens of repository managers. Retrieved April 21, 2025, from https://repositorium.sdum.uminho.pt/bitstream/1822/60527/3/OpenAIRE_OR2019_workshop_2nd_all.pdf
- RDA FAIR Data Maturity Model WG. (2020). FAIR Data Maturity Model: Specification and
-

- Guidelines. Retrieved April 21, 2025, from <http://www.rd-alliance.org/groups/fair-data-maturity-model-wg>
- Rhee, G., Yurb, P., & Ryoo, S. Y. (2020). Performance Measurement Model for Open Big Data Platform. *Knowledge Management Research*, 21(4), 243-263. <https://doi.org/10.15813/kmr.2020.21.4.013>
- Saaty, T. L. (1980). *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. Virginia: McGraw-Hill International Book Company.
- Shin, J. (2021). Data quality verification method for artificial intelligence learning. *Journal of the Electronic Engineering Society*, 48(7), 28-34.
- Singapore Economic Development Board. (2020). The Smart Industry Readiness Index. Retrieved April 21, 2025, from <https://www.edb.gov.sg/en/about-edb/media-releases-publications/advanced-manufacturing-release.html>
- Stuart, D., Baynes, G., Hrynaszkiewicz, I., Allin, K., Penny, D., Lucraft, M., & Astell, M. (2018). Whitepaper: Practical challenges for researchers in data sharing (Version 1). figshare. <https://doi.org/10.6084/m9.figshare.5975011.v1>
- Telecommunications Technology Association. (2021). *Data quality management requirements for supervised learning (TTAK.KO-10.1339:2021)*.
- Yoo, S. (2012). A study on evaluation model of business process management systems based on analytical hierarchy process. *Management & Information Systems Review*, 31(4), 433-444. <https://doi.org/10.29214/damis.2012.31.4.018>
- Yu, J. (2012). *Professor Jong-pil Yu's concept and understanding of structural equations*. Seoul: Hannarae Publishing House.

[About the authors]

Seonghun Kim is an Affiliate Professor of Sungkyunkwan University iSchool: LIS and Data Science, Seoul, The Republic of Korea. He teaches courses on web databases, system design and information retrieval. His main research interests are deep learning and LLM-based recommendation systems, and he is researching academic paper recommendation systems with research support from the National Research Foundation of Korea.

Jinho Park is an Assistant Professor in the Knowledge, Information and Culture Track at Hansung University, South Korea. He teaches courses on databases, data analysis and information retrieval. He is a board member of the Korean Library and Information Society, the Korean Society of Records Management, and the Korean Library Association, and an expert member of ISO/TC46 Korea. He has led the informatics department of the National Digital Library of Korea and the Arts Information Centre of the Korea National University of Arts. His main research interest is open data. His research focuses on building new information ecosystems and management and service systems through data openness, including open data quality assessment.
