J. Internet Comput. Serv.
ISSN 1598-0170 (Print) / ISSN 2287-1136 (Online)
http://www.jics.or.kr

Suggestion for an ISO 25010 quality model encompassing AI-based software☆

Seung-Hee Kim[1*]

ABSTRACT

This study developed a novel ISO/IEC 25010 quality model for the quality management of artificial intelligence (AI)-based software by using quality characteristics classification card (QCCC) quality models. We used AI models to add, modify, and restructure AI quality attributes for the product quality model and the quality-in-use model of the ISO/IEC 25010 quality model to derive a novel ISO/IEC 25010 quality model. By integrating quality standards derived from various AI-related models, we enhanced the accuracy of the derived model. The product quality model included 10 main quality and 45 subquality attributes, and the quality-in-use model included 10 main quality and 28 subquality attributes. In AI-based models, the quality-in-use model was found to require modifications. The results revealed the direction of improvement of the AI-compatible software quality model and the possibilities for potential standardization and conflict resolution. This study presents the direction for standardization reviews on reorganizing the quality attributes, concepts of attributes, and relationships so that they can be applied to AI software while maintaining the framework of the currently defined software quality model. The results can serve as criteria for the quality management of AI-based software and can also contribute to research on quality models for AI-based software.

☞ keyword : artificial intelligence quality, ISO/IEC 25000, product quality, quality-in-use, software quality

1. Introduction

Artificial intelligence (AI)-based intelligent products are typically implemented through data-learning-based modeling and are equipped with cognitive and decision-making capabilities to achieve desired goals [1]. The European Union's Artificial Intelligence Act requires a quality management system to ensure that AI systems are operated as intended [2]. According to the act, the quality management system includes the techniques used in AI system design, design control, and design verification, techniques used in development, quality control, and quality assurance, the inspection and testing procedures to be performed before, during,and after system development, and the testing frequency [2,3]. Moreover, the quality management system should be implemented through documents covering the following areas: technical specifications, including standards, means used to ensure compliance with statutory requirements in case standards do not apply, systems and procedures for data management, including data acquisition, collection, analysis, labeling, storage, filtering, aggregation, and data mining, risk management system, setting up, implementing and maintaining a post-sales monitoring system, and information to be communicated to national agencies and certification bodies that provide or support access to data [2,4]. Generally, AI-based systems are widely used in quality control elements such as flexibility, adaptability, autonomy, evolution, bias, transparency/ interpretability/ explainability, complexity, and nondetermination [5-9]. These parameters are used as checklists to identify project and system risks when setting up an AI system quality test plan.

The International Organization for Standardization and the International Electrotechnical Commission 25000 (ISO/IEC 25000) [2][10] is the most representative international standard for the quality assessment of general software (SW) and includes five standard areas of software quality control, quality model, quality measurement, quality requirements, and quality assessment [2,11] guidelines. In particular, the software quality model, defined in 25010, is categorized into product quality and quality-in-use. Functional suitability, performance efficiency, compatibility, usability, reliability, security, maintainability, and portability are primary quality attributes [10]. Functional

1 Dept. of IT Convergence Software Engineering, Korea University of Technology & Education (KOREATECH)., Cheonan, 31253, Korea
* Corresponding author: Seung-Hee Kim (sh.kim@koreatech.ac.kr)
[Received 09 July 2024, Reviewed 27 August 2024 Accepted 24 September 2024]
☆ This paper was supported by the Education and Research Promotion Program of KOREATECH in 2023

http://dx.doi.org/10.7472/jksii.2024.25.5.67

suitability, a primary quality attribute, is the standard for functions required of a system, whereas the remaining seven main quality attributes are the standard of a quality model for nonfunctions. Each of these main quality attributes has 31 sub-quality attributes [12].

Quality in use is a measurement standard of the quality of the product used by the person and the service and consists of effectiveness, efficiency, safety, satisfaction, and context coverage. Furthermore, high, average, and low weights were given to effectiveness, satisfaction, and safety, respectively.

However, software quality management for AI-based intelligent products remains a challenge in the software (SW) quality model of ISO/IEC 25010. The implementation and maintenance of AI-based software can be attributed to the unique quality characteristics for AI platforms, AI models, AI-related software, and AI data that are not presented in the ISO/IEC 25010 model because of the unique AI system characteristics. Federer and Ramler [13] derived the unique characteristics of AI-based systems. They highlighted several challenges in AI systems, including the complexity and lack of transparency of AI models, unclear specifications and requirements, and the need for comprehensive validation data and test inputs. Moreover, the need to establish reliable benchmarks for outcomes, assess accuracy and correctness, address non-functional aspects, and adapt to the self-learning nature and evolving conditions of these systems was indicated.

Therefore, EU, IEEE, NIST, IECD, and NESCO [14] have defined standards and guide-lines for the development and acceptance of AI systems. Conventional software testing approaches are still required when testing an AI-based system. However, AI-based systems include numerous special attributes that can make additional testing necessary than for conventional software systems [14].

Therefore, the development of a novel international quality management standard that covers all AI-based intelligent software and systems is crucial. For AI, AI platform, and AI systems, Kharchenko et al. [14] developed a quality model consisting of a total of 46 attributes that can be used in the creation, implementation, and evaluation of AI systems, as well as the regulation of technology and tool development for AI system standardization.

This study is focused on the differences between the conventional ISO/IEC 25010 quality model and the AI quality model derived from the study in [14] and provides direction for improving the novel ISO/IEC 25010 quality model that can be used to the AI model. The results of this study can be used as a reference for the development of quality models of novel international standards that can cover both AI-based intelligent products and general software and provide a theoretical basis for developing new standards.

2. Theoretical Background

2.1 Trends of ISO/IEC 25000 Standardization based on AI

According to the Joint Technical Committee ISO/IEC JTC 1, Information Technology, Subcommittee SC 7, Software and System Engineering [15,16], a standard revision is on-going because AI-based systems conflict or are inconsistent with the conceptual definitions of the main and sub-quality attributes of the ISO/IEC 25010 quality model standard (Table 1) [10,15].

The ISO/IEC 25010 quality model consists of eight main quality attributes and 31 sub-quality attributes [17,18]. Functionality suitability, a main quality attribute [6,14] as-sociated with functional requirements, is composed of functional completeness, correctness, appropriateness [17]. The remaining quality attributes are nonfunctional requirements [6]. The other main quality attributes are composed of the following sub-quality attributes [6,17,18]. Performance efficiency consists of time behavior, resource utilization, and capacity, and compatibility consists of coexistence and interoperability. Usability consists of appropriateness recognizability, learnability, operability, user error protection, user interface aesthetics, and accessibility. Reliability consists of maturity, availability, fault tolerance, and recoverability. Security consists of confidentiality, integrity, nonrepudiation, accountability, and authenticity. Maintainability consists of modularity, reusability, analyzability, modifiability, and testability. Portability consists of adaptability, installability, and replaceability.

The quality metrics applied to the test guidelines have been developed in accordance with ISO/IEC 25059. The key test items to be applied to AI in accordance with ISO/IEC TR 29119-11 [6], which is included in guidelines on the testing of AI-based systems, is applied to quality models. These items can be distinguished between the added quality items and quality items in which the concept should be adjusted. First, we investigate the quality items that are added. Functional adaptability is addedto sub-quality attributes included in the functional suitability, which is a main quality attribute of ISO/IEC 25010. This attribute is defined as the extent to which an AI system accurately [19] obtains information from the results of data or previous operations and can use that information for future predictions [6,19]. User controllability and transparency sub-quality attributes are then added to the usability main quality attribute of ISO/IEC 25010. User controllability is defined as the extent to which users can intervene in the operation of an AI system [19] in an appropriate manner [6], and transparency is defined as the extent to which appropriate information about the AI system is communicated to stakeholders [6,19].

For reliability, one of the main quality attributes of ISO/IEC 25010 [20], robustness is added as a sub-quality attribute, which is defined as the extent to which an AI system can maintain [19] a performance level in any situation [6]. Furthermore, intervenability can be added to the security main quality attribute, which is defined as the extent to which an operator can intervene at the right time in the operation of an AI system to avoid damage or risk.

Functional correctness is being reviewed in the ISO/IEC 25010[20] quality standard and is included in the functional suitability main quality attribute. Originally, functional correctness is defined [21] as the degree to which a product or system provides the correct results with the needed degree of precision [6,21]. However, among AI systems, especially in the case of machine learning methods, determining whether functional correctness is accurate and correct in all observed situations is difficult [6].

Materials and Methods should be described with sufficient details to allow others to replicate and build on the published results. Please note that the publication of your manuscript implicates that you must make all materials, data, computer code, and protocols associated with the publication available to readers. Please disclose at the submission stage any restrictions on the availability of materials or information. New methods and protocols should be described in detail while well-established methods can be briefly described and appropriately cited.

Research manuscripts reporting large datasets that are deposited in a publicly avail-able database should specify where the data have been deposited and provide the relevant accession numbers. If the accession numbers have not yet been obtained at the time of submission, please state that they will be provided during review. They must be provided prior to publication.

Interventionary studies involving animals or humans, and other studies that require ethical approval, must list the authority that provided approval and the corresponding ethical approval code.

2.2 Studies on AI Quality Attributes

This section examines studies on various topics related to AI-based quality, including the European Union Charter on Artificial Intelligence and the Japanese QA4AI Consortium's ML-based AI System Quality Assurance Guidelines. The results of the studies were summarized by subclassifying them into AI product and service quality assessment, an AI-based quality management system case studies, and AI ethics criteria.

First, studies on the quality assessment of AI products and services were reviewed. Kharchenko et al. [14] defined AI, AI platform (AIP), and AI system (AIS) and presented a hierarchy of quality models and quality attributes for each concept. Felderer & Ramler [13] identified the characteristics and challenges of AISs through research on the exact characteristics of the AI system required for the quality assurance of AI-based intelligent systems. By identifying the approaches and challenges associated with AI-based systems in software engineering for AI-based systems, Silverio et al. [1] demonstrated that reliability and safety are the most studied quality attributes.

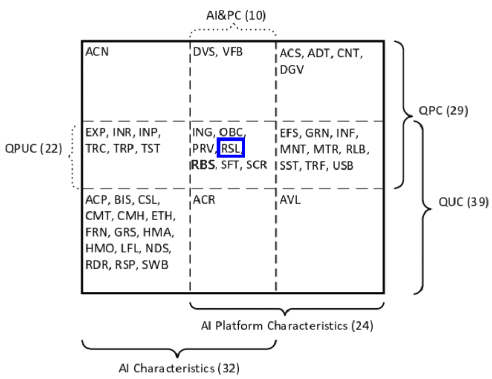
Jayakumar et al. [22] developed quality assessment criteria that systematically support high accuracy of diagnostic analysis and can be applied to the AI-based health diagnostics system. Manziuk et al. conducted studies to develop a reliable AI orthodontic model based on reliability and ontology concepts [23] and a standardized development standard model of formal approaches for the rapid deployment and application.

Next, studies in the AI users and service quality section present the direction of the evolution of the network as an intelligent autonomous infrastructure by studying the development of network technology and technology trends [24,25], derived preliminary quality assessment items related to AI services and presented quality evaluation items including appropriate AI services through the Delphi technique [26], and developed a multidimensional quality assessment system supporting customer center service quality assessments using AI [27]. Chen et al. [28] defined the standard service quality levels and the dimensions of service quality in various AI environments and proposed a multi-stage quality classification and measurement model for mixed application based on AI Chatbot service cases. Baek et al. [29] investigated the necessity of quality assessment attributes of AI services through previous research analysis and presented seven main quality [30] attributes and 24 detailed quality assessment items [30].

Many countries, enterprises, and organizations have adopted ethical standards for artificial intelligence, standardization, and AI regulatory frameworks [31]. From risk management perspective, the type of regulation, reliability, transparency, and accountability are representative quality items [32]. From data management perspective, privacy violations, security, as well as the impact on health, security, and rights are representative items.

2.3 Kharchenko's QCCC Quality Model

From data management perspective Kharchenko et al. [14] categorized the quality of an AIS into AI quality and the AIP quality and classified the attributes of product quality and quality-in-use. They defined AI quality as a specific attribute to which the unique set of attributes of an object satisfies the requirements and categorized AIP quality into the quality of the software and hardware platforms implementing AI. This study summarized

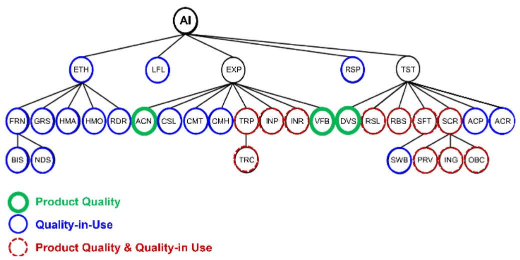


(Figure 1) Classification map of AI quality characteristics [14].

the findings of Kharchenko et al. through the classification map of AI quality characteristics [14] (Figure 1)

Figure 1 displays the quality characteristics classification card (QCCC) [14] quality system chart. The columns correspond to AI, AIP, and AI&PC. In QCCC columns, AI&PC refers to the common part of AI characteristics and AIP characteristics [14]. Furthermore, the rows correspond to quality assessment classification, representing quality-in-use (QUC), product quality characteristic (QPC), and characteristics that include both types of quality (QPUC) [14].

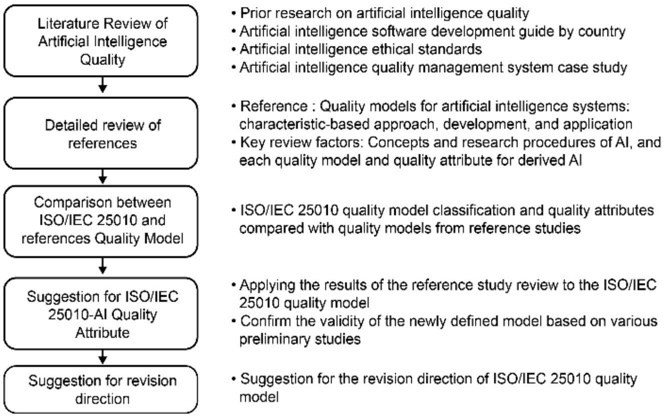
The quality attributes of Level 1, the candidates for the quality model of AI and AI platforms, are mapped in the model. For example, RSL is a quality attribute that is suitable for both AI quality and AIP quality assessment with the highest level of resiliency quality. Next, the quality attributes of the AI top quality model and the AIP top quality model were derived comprehensively and combined to complete the AIS quality model. Because this method is a study on the applicability of software quality models, the AI quality model is used as displayed in Figure 2.



(Figure 2) Graphical representation of the AI quality model of QCCC

3. Research Procedure and Method

Figure 3 displays the research procedure of this study. Based on research on AI platforms, AI services, AISs, a study by Vyacheslav et al. [14], and the results of the analysis of AI quality attributes, the quality attributes included in the conventional ISO/IEC 25010 quality model were compared with those of the QCCC quality model [33]. Improvements to the novel ISO/IEC 25010 quality model were suggested by evaluating the results of the review and other studies related to AI quality. A quality model of ISO/IEC 25010 that encompasses



(Figure 3) Procedure for using the ISO/IEC 25010 quality model for enhancing AISs adaptation research

AI software was proposed. Finally, the features and implications of the proposed model are discussed.

4. Development of a Quality Model for AI System Applications

4.1 Development of an ISO/IEC 25010-AI model

As the first step in the development of a quality model for AI system applications, the model's attribute names and defined attributes for the AI quality model of the QCCC quality model and the product quality and quality-in-use models of the ISO/IEC 25010 were compared. In the QCCC quality model, attributes of both 'the product quality' and 'quality-in-use' models (Figure 3) are used. Furthermore, some of the attributes defined in the QCCC quality model [34] have the same attribute names as the quality attributes defined in the ISO/IEC 25010, but in some attribute names, definitions do not match exactly. In this procedure, the differences between the two models are analyzed through comparison, and items are then added to the ISO/IEC 25010 to enable a comprehensive application of AI quality, or if the same quality attribute exists, the concept is verified and the redefinition requirement is checked.

First, the result of comparative analysis of the definitions of the QCCC's AI quality model with the ISO/IEC 25010 product quality model [10] summarized in Table 1 revealed the necessity of the addition of the explainability (EXP) and trustworthiness (TST) main quality attributes and four sub-quality attributes per main quality attribute. However, interoperability (INP) derived from EXP's sub-quality attributes was excluded because it is already included in the ISO/IEC 25010. By contrast, EXP's Accountability (ACN) exists as a sub-quality attribute of the Security main quality attribute of the ISO/IEC 25010 [35], and the concepts vary and should be replaced by other terms that can express the concept. The definition of the concept of accountability in the security main quality attribute of the ISO/IEC 25010 [35] then the extent to which it is traceable in terms of security. By contrast, the QCCC's conceptual definition of the ACN, the sub-quality attribute of EXP, is interpreted as the ability of AI to report/explain work results in a transparent manner in a defined format. Thus, in this study, traceability was determined to be more suitable than the accountability included in the security main quality attribute of the ISO/IEC 25010 [35]. This phenomenon was reflected in the mapping model.

In addition, security (SCR) was derived as a sub-quality attribute of TST in the QCCC, but it is defined as a main quality attribute in the ISO/IEC 25010. Thus, mapping privacy (PRV), integrity (ING), and objectivity (OBC), which had been derived from the QCCC as the detail attributes of SCR based on the ISO/IEC 25010, as sub-quality attributes is necessary. Furthermore, integrity should be redefined by synthesizing with the existing sub-quality attribute so that it can be applied to AI and AIPs.

Next, the QCCC quality model and the ISO/IEC 25010 quality-in-use model were compared and mapped. As presented in Table 2, five main quality attributes and the security sub-quality attribute should be added.

(Table 1) Test items and verification details

Main quality attribute	Sub/Detailed Quality Attributes		Overview for applying quality attributes
Explain-ability (EXP)	Interactivity (INR)		Additional reflection of the main quality and sub-quality attributes
	Transparency (TRP)		Additional reflection of the main quality and sub-quality attributes
	Verifiability (VFB)		Additional reflection of the main quality and sub-quality attributes
	Accountability (ACN)		The attribute "accountability" has been redefined as "traceability" in this study to align more closely with its conceptual definition according to ISO/IEC 25010
Trustworthiness (TST)	Diversity (DVS)		Additional reflection of the main quality and sub-quality attributes
	Resiliency (RSL)		Additional reflection of the main quality and sub-quality attributes
	Robustness (RBS)		Additional reflection of the main quality and sub-quality attributes
	Safety (SFT)		Additional reflection of the main quality and sub-quality attributes
	Security (SCR)	Privacy (PRV)	All sub-quality attributes of Trustworthiness are mapped to the main quality attributes of ISO/IEC 25010, and the detailed quality attributes are added as sub-quality attributes. However, it is redefined by reflecting the definition of the attribute defined by QCCC or by combining it with the existing concept.
		Integrity (ING)	
		Objectivity (OBC)	

Finally, the final model of the ISO/IEC 25010 that could encompass AI was confirmed and quality items [6] for testing AI-based systems was reflected in the results according to the ISO/ IEC 29119. Therefore, functional adaptability was added as a sub-quality attribute to functional suitability, a main quality attribute of the ISO/IEC 25010 [35] quality model, user controllability and transparency were added to usability, robustness was added to reliability, and intervenability was added to security.

(Table 2) Analysis results obtained through the mapping of the QCCC's model to the ISO/IEC 25010 quality-in-use model.

Main quality attribute	Sub/detailed quality attributes		Overview for applying quality attributes
Ethics (ETH)	Fairness (FRN), Graspability (GRS), Human agency (HMA), Human oversight (HMO), Redress (RDR)		Additional reflection of the main quality and sub-quality attributes
Explainability (EXP)	Causability (CSL), Completeness (CMT), Comprehensibility (CMH), Interactivity (INR), Interpretability (INP), Transparency (TRP)		Additional reflection of the main quality and sub-quality attributes
Trustworthiness (TST)	Acceptability (ACP), Resiliency (RSL), Robustness (RBS), Safety (SFT)		Additional reflection of the main quality and sub-quality attributes
	Security (SCR)	Privacy (PRV), Integrity (ING), Objectivity (OBC)	SCR sub-quality attributes are mapped to the Security main quality attributes of ISO/IEC 25010, and the security attribute concept is newly defined to include all of the concepts of Privacy (PRV), Integrity (ING), and Objectivity (OBC) of detailed quality attributes
Lawfulness (LFL)	-	-	Additional reflection of the main quality
Responsibility (RSP)	-	-	Additional reflection of the main quality

These attributes were included as the attributes of the quality model because adopting them as requirements was desirable in the ISO/IEC international standard in terms of consistency.

Tables 3 and 4 present the AI-based product quality attribute model and the AI-based quality-in-use attribute model, respectively. The definitions in these tables followed those of the previous studies as much as possible, maintaining the quality attribute concepts as defined by the ISO/IEC 25010 and QCCC. However, duplicate definition or redefined quality attributes were redefined by combining the concepts of the references and these new definitions were described in the 'Definition' column. The 'Ref.' column presents the previous studies referred to derive quality attributes and define existing and new concepts.

〈Table 3〉 Results of the initial development of the ISO/IEC 25010 product quality model 〔36〕 encompassing AI

Quality attributes		Definition of quality attributes	Ref.
Main	Sub		
Functional suitability	Functional completeness	Same as ISO/IEC 25010 definition	〔17〕
	Functional correctness	Degree to which the functions or performance provide the correct results with the needed quality metric	〔6〕, 〔14〕, 〔37〕
	Functional appropriateness	Same as ISO/IEC 25010 definition	〔17〕
	Functional Adaptability	Same as ISO/IEC 25010 definition	〔38〕
Performance efficiency	Time-behavior	Same as ISO/IEC 25010 definition	〔17〕
	Resource utilization		
	Capacity		
Compatibility	Co-existence	Same as ISO/IEC 25010 definition	〔17〕,〔18〕
	Interoperability		
Usability	Appropriateness recognizability	Same as ISO/IEC 25010 definition	〔17〕,〔18〕
	Learnability		
	Operability		
	User error protection		
	User interface aesthetics		
	Accessibility		
	User controllability	Redefine by combining existing definitions	〔6〕, 〔14〕, 〔38〕
	Transparency	Redefine by combining existing definitions	〔6〕, 〔14〕, 〔38〕
Reliability	Maturity	Same as ISO/IEC 25010 definition	〔17〕,〔18〕
	Availability		
	Fault tolerance		
	Recoverability		
	Robustness	Redefine by combining existing definitions	〔6〕, 〔14〕, 〔38〕
Security	Confidentiality	Same as ISO/IEC 25010 definition	〔17〕,〔18〕
	Integrity	Redefine by combining existing definitions	〔6〕, 〔14〕, 〔17〕,〔18〕
	Nonrepudiation	Same as ISO/IEC 25010 definition	〔17〕,〔18〕
	*Traceability	Change by combining existing name, definitions	〔14〕, 〔17〕, 〔18〕
	Authenticity	Same as ISO/IEC 25010 definition	〔17〕,〔18〕
	Intervenability	Change by combining existing name, definitions	〔6〕, 〔38〕
	Objectivity	Same as QCCC definition	〔14〕
	Privacy	Same as QCCC definition	〔14〕
Maintainability	Modularity	Same as ISO/IEC 25010 definition	〔17〕,〔18〕
	Reusability		
	Analyzability		
	Modifiability		
	Testability		
Portability	Adaptability	Redefine by combining existing definitions	〔14〕, 〔17〕,〔18〕
	Install-ability	Same as ISO/IEC 25010 concept	〔17〕,〔18〕
	Replaceability		
Explainability	Interactivity	Same as QCCC definition	〔14〕
	Transparency		〔14〕
	Verifiability		〔14〕
	Accountability		〔14〕
	Interpretability		〔14〕
Trustworthiness	Diversity	Same as QCCC definition	〔14〕
	Resiliency		〔14〕
	Safety		〔14〕

* Traceability in security is defined in the ISO/IEC 25010 standard as the sub-quality attribute accountability. However, this definition is the same as the sub-quality attribute name (accountability) included in the EXP main quality attribute of the QCCC quality model defined to cover AI. Thus, based on the concept of the term as defined in the ISO/IEC 25010 and the QCCC model, traceability is suited to the quality attributes in the ISO/IEC 25010. Therefore, we made the change

〈Table 4〉 Results of the initial development of the ISO/IEC 25010-AI quality-in-use model

Quality attributes		Definition of quality attributes	Ref.
Main	Sub		
Effectiveness	-	Same as ISO/IEC 25010 definition	〔18〕
Efficiency	-	Same as ISO/IEC 25010 definition	〔18〕
Satisfaction	Usefulness	Same as ISO/IEC 25010 definition	〔18〕
	Trust		
	Pleasure		
	Comfort		

Quality attributes		Definition of quality attributes	Ref.
Main	Sub		
Freedom form Risk	Economic Risk Mitigation	Same as ISO/IEC 25010 efinition	[18]
	Health and Safety Risk Mitigation		
	Environmental Risk Mitigation		
	Security	Redefine by combining existing unique definitions & QCCC definitions	[14]
Context Coverage	Context Completeness	Same as ISO/IEC 25010 definition	[18]
	Flexibility		
Ethics	Fairness	Same as QCCC definition	[14]
	Graspability		
	Human agency		
	Human oversight		
	Redress		
Explainability	Causability	Same as QCCC definition	[14]
	Completeness		
	Comprehensibility		
	Interactivity		
	Interpretability		
	Transparency		
Trustworthiness	Acceptability	Same as QCCC definition	[14]
	Accuracy		
	Resiliency		
	Robustness		
	Safety		
Lawfulness	-	Same as QCCC definition	[14]
Responsibility	-	Same as QCCC concept	[14]

4.2 Validation and supplementation through the coverage of the ISO/IEC 25010 quality model that applies AI quality attributes

This procedure validates the ISO/IEC 25010 product quality model and quality-in-use model that reflect the AI quality attributes. Therefore, the quality attributes of references (a) in Table 5 were compared with those derived from studies on AI software quality-related topics between 2018 and 2023, as a preliminary study, with the exception of [14]. This result was designed as a method to verify the appropriateness and feasibility of the ISO/IEC 25010 model that encompasses AI-based software by comparing the random quality attributes column (a) in Table 5 with the mapping models in Tables 3 and 4.

The presence of quality items was determined to verify the feasibility of the models. To derive quantitative measurement results, the coverage ratio of the quality attributes was measured as a metric. The quality attribute coverage is calculated as a ratio of the number of the quality attributes (a) in Tables 3 and 4, the result of Step 1, to the total number (n) of quality attributes derived from the references [39], as presented in Equation 1.

$$\text{Coverage Ratio (CR)} = (\text{exist\_count}(a)/n)*100 \quad (\text{CR} <= 100) \qquad (1)$$

⟨Table 5⟩ Validation results and coverage of the ISO/IEC 25010 quality model to encompass AI through the application of quality attributes defined or applied in AI-related studies over the last five years.

Source classification	Ref.	Metrics for applying quality attributes to validate the derived quality model (a)	Result of mapping attributes (a) to Tables 3 and 4 (b)	Quality attribute coverage ratio (CR)
Case studies of AI-based system development	[40]	Data loss (reliability)	Reliability-Recoverability	100%
		Data collection quality (data transparency)	Security-Objectivity	
		Traffic prediction algorithm (prediction performance)	Functional Suitability-Functional Correctness	
	[41]	Reading welding defects	Explainability-Transparency, Usability, Transparency	100%
		Performance of visualization algorithm	Explainability-Functional Correctness	
AI ethics	[32]	Perspective of risk control regulation type: reliability, transparency, accountability	Trustworthiness-Safety, Explainability-Transparency, Responsibility	100%
		Data management: availability, quantity, suitability	Reliability-Availability & Robustness, Functional Suitability-Functional Adaptability, Trustworthiness-Diversity Security-Objectivity	100%
		Individual's health/safety/basic rights	Trustworthiness-Safety	
		Privacy violation	Security-Integrity & Privacy	

Source classification	Ref.	Metrics for applying quality attributes to validate the derived quality model (a)	Result of mapping attributes (a) to Tables 3 and 4 (b)	Quality attribute coverage ratio (CR)
		Data pseudonymization and access control	Security-Confidentiality & Integrity	
		Enhancement of security and resilience for technical or physical disasters	Trustworthiness-Safety & Resiliency, Security-Confidentiality, Reliability-Recoverability & Robustness	
Product quality	[15]	Prediction performance (Precision)	Functional Suitability-Functional Correctness	100%
		Performance of data (ground information) and outlier detection models	Functional Suitability-Functional Adaptability	
	[2]	Prediction of test days	Functional Suitability-Functional Correctness & Functional Adaptability	100%
		Failure data quality	Trustworthiness-Diversity, Security-Objectivity	
		Time spent on testing	Performance Efficiency-Time-behavior " Resource utilization	
	[42]	Software maintainability (software complexity (Macabe's Cyclomatic Complexity) coverage)	Maintainability Modularity & Testability	100%
	[43]	Metadata verification, large DB processing capacity, data life cycle management	Maintainability-Testability, Explainability-Verifiability, Performance Efficiency-Capacity, Maintainability Reusability	100%
	[44]	Data integrity	Security-Confidentiality & Integrity	100%
		Model robustness	Reliability-Robustness Trustworthiness-Robustness	
		System quality	Reliability-Availability & Robustness	
		Process agility	Explainability-Interactivity, Maintainability-Analyzability, Functional Suitability-Functional Appropriateness	
		Customer expectation	Reliability-Maturity & Availability Explainability-Interactivity	
	[45]	Noise data	Trustworthiness-Diversity, Security-Objectivity & Intervenability	100%
		Detection of initial pipeline data errors, validity of data verification process, validity based on inferred schema, validity of unstructured data verification feature	Explainability-Verifiability & Intervenability, Functional Suitability-Functional Appropriateness	
		Outlier detection techniques and calculations	Functional Suitability-Functional Correctness	
Use & services quality	[24]	Tracking management of QoS-related events	Explainability-Transparency	100%
	[25]	Accurate detection of current events, automatic mitigation, immediate processing level	Explainability-Completeness & Comprehensibility	100%
		Reliable prediction of QoS-related events	Trustworthiness-Acceptability & Resiliency	
	[26] [46]	Real-time: quick reaction, interaction, simultaneity	Explainability-Interpretability & Interactivity	71.4%
		Personalization: customer characteristics analysis, continuous monitoring	N/A	
		Expertise: accurate response, high quality, latest information [46]	N/A	
		Diversity: service connection, need recognition, complex processing	Context Coverage-Context Completeness & Flexibility, Explainability-Causability	

Source classification	Ref.	Metrics for applying quality attributes to validate the derived quality model (a)	Result of mapping attributes (a) to Tables 3 and 4 (b)	Quality attribute coverage ratio (CR)
		Convenience: autonomy, convenience of use, cost effectiveness, enjoyment, comfort	Satisfaction-Pleasure & Comfort Efficiency	
		Reliability: service stability, personal information protection, perfect [46] performance, problem response, privacy [46]	Trustworthiness-Safety, Freedom form Risk-Security	
		Spatiotemporality: no restriction on time, no restriction on place	Context Coverage-Flexibility & Context Completeness	
	[27]	Interaction quality	Explainability-Interactivity & Interpretability	60%
		System quality: systematicity (rationality), convenience, diversity, autonomy, simultaneity	N/A	
		Result quality: accuracy	N/A	
		Result quality: rapidity	Trustworthiness-Robustness	
		Result quality: reliability	Satisfaction-Trust	

The newly defined ISO/IEC 25010-AI was validated by extracting a variety of metrics measured in relation to the quality of AI in AI-related studies. Therefore, the coverage ratio for the overall product quality model was 100% and that of the quality-in-use model was 82.85%. The cause of the low coverage of the derived quality-in-use model was analyzed. First, a quality attribute that can measure the level of personalization of various types of users is yet to be established. Second, the existing general software did not encounter situations in which most users had to maintain or manage the functionality or performance. Therefore, satisfaction was not related to product performance. However, because AI-based software deal with product performance in terms of user utilization, a model is required for measuring system quality attributes that can comprehensively measure the expertise and specificity of systems, such as rationality, diversity, and simultaneity. Thus, to address this problem, specialization and personalization were added as the sub-quality attributes of the satisfaction main quality attribute of the quality-in-use model [47].

4.3 Definition of the ISO/IEC 25010 quality model that applies AI quality attributes

Tables 6 and 7 present the finalized product quality model and quality-in-use model of ISO/IEC 25010 for comprehensively supporting AI quality control. The newly defined concepts of quality attributes reflect the results of the present study, and all the remaining quality attributes have been quoted as defined in the reference literature, such as ISO/IEC 25010 [6,14]. However, their applicability and usability as a quality model were enhanced by expressing them comprehensively through software, without prescribing AI software. The concept definition way (CDW) column defines quality attributes. Here, R denotes the concepts of attributes quoted from the definitions in the references, $\hat{R}$ denotes the concepts quoted after partially changing them, and S denotes the concepts defined in this study. The underline in the definition indicates the changed or added parts.

(Table 6) Definition of the proposed ISO/IEC 25010 product quality model encompassing AI

Quality characteristics		Definition	CDW
Main	Sub		
Functional Suitability	Functional Completeness	Degree to which the set of functions covers all the specified tasks and user objectives [10,17].	R
	Functional Correctness	Degree to which the functions, data, or performance provide the correct results with the needed quality metric [10,17,19].	$\hat{R}$
	Functional Appropriateness	Degree to which the functions or data facilitate the accomplishment of specified tasks and objectives [10,17,19].	$\hat{R}$
	Functional Adaptability	Degree to which software and used data can accurately acquire information from previous data, or the result of previous actions, and use that information in future predictions [38].	$\hat{R}$

Quality characteristics		Definition	CDW
Main	Sub		
Performance Efficiency	Time-behavior	Degree to which the response and processing times and throughput rates of a product or system, when performing its functions, satisfy requirements (10,17,19).	R
	Resource Utilization	Degree to which the specification, amounts, and types of resources used by a product or system, when performing its functions, satisfy requirements(10,17,19).	R̂
	Capacity	Degree to which the maximum limits of the product or system, parameter satisfy requirements (10,17,19).	R
Compatibility	Co-existence	Degree to which a product can perform its required functions efficiently (10) while sharing a common environment and resources with other products(10), without detrimental effect on any other product(10,18,19).	R
	Interoperability	Degree to which two or more systems, products, or components can exchange information and use the information that has been exchanged (10,17,18,19,48).	R
Usability	Appropriateness recognizability	Degree to which users can recognize whether (10) a data, a feature module, product, or system is appropriate for their needs (10,17-19).	R̂
	Learnability	Degree to which a feature module, product, or system enables the user to learn how to use it with effectiveness, efficiency (10,17-19).	R̂
	Operability	Degree to which a feature module, product or system is easy to operate, control, and appropriate to use (10,17-19).	R̂
	User error protection	Degree to which a feature module, a product, or system protects users against making errors(17-19).	R̂
	User interface aesthetics	Degree to which a user interface enables pleasing and satisfying interaction for the user (10,17-19).	R

Quality characteristics		Definition	CDW
Main	Sub		
	Accessibility	Degree to which a feature module, product, or system can be used by people with the widest range of characteristics and capabilities to achieve a specified goal in a specified context of use (17-19).	R̂
	User Controllability	Degree to the extent to which IT governance frameworks empower users to oversee, manage, and make informed interventions in data, AI systems, and processes, ensuring these technologies perform as expected.	S
	Transparency	Degree to which software enables the illustration, examination, or reproduction of models, individual components, and decision-making algorithms, ensuring relevant information is effectively conveyed to stakeholders.	S
Reliability	Maturity	Degree to which a system, product, or component (10) satisfies the requirements for reliability under(19) standard operational conditions, with the capability as determined by its compliance with reliability standards specified by the client (17-19).	R̂
	Availability	Degree to which a feature module, product, or system is operational and accessible when required for use (17,18).	R̂
	Fault tolerance	Degree to which a system, product, or component operates as intended despite the presence of hardware or software faults (10,19,43,44).	R
	Recoverability	Degree to which, in the event of an interruption or a failure, a product or system can recover data directly affected and re-establish the desired state of the system (10,17-19).	R
	Robustness	Level/degree of functionality that allows the system to operate by maintaining performance levels over a range of input data and operating conditions, and reliably enter a shutdown state when data and conditions exceed specified limits.	S

Quality characteristics		Definition	CDW
Main	Sub		
Security	Confidentiality	Degree to which the product or feature module ensures that data are accessible only to those authorized to have access (17-19).	R̂
	Integrity	Degree to which a system, product or component, feature module, platform prevents unauthorized access to, or modification of computer programs, AI or data (18,19).	R̂
	Nonrepudiation	Degree to which actions or events can be proven to (10) have occurred so that the events (19) or actions cannot be repudiated (10,17-19).	R
	*Traceability	Degree to which the actions of ensuring relevant information, an algorithm, product, feature module, system can be traced uniquely to the entity (14).	S
	Authenticity	Degree to which the identity of a subject or resource can be proved to be the one claimed (10,17,19)	R
	Intervenability	Degree to which an operator can intervene in the operation of an AI or system in a timely manner to avoid damage or danger (19,38).	R
	Objectivity	Ability of AI, product, system, and platforms to prevent the use of corrupted or falsified data (14).	R̂
	Privacy	Degree of features of software, data, and platform that guarantee the right to retain personal information according to user requirements (25).	R̂
Maintainability	Modularity	Degree to which a system or computer program is composed of discrete components such that a change to one component has minimal effect on other components (17-19).	R
	Reusability	Degree to which an asset can be used in more than one system, or in building other assets (17-19).	R
	Analyzability	Degree of effectiveness and efficiency with which it is possible to assess (10) the effect on a product (10) or system of an intended change to one or more of its parts (10), or to	R
		diagnose a product for deficiencies (10) or causes of failures, or to identify parts to be modified (10,17-19).	
	Modifiability	Degree to which a product or system(10) can be effectively and efficiently modified(19) without introducing defects or degrading existing product quality (10,17-19).	R
	Testability	Degree of effectiveness and efficiency with which test criteria can be established (10) for a feature module, system, product or component and tests can be performed to determine whether those criteria have been satisfied (17-19).	R̂
Portability	Adaptability	Ability to the extent to which a product or system can be effectively and efficiently modified to adapt to various or evolving hardware, software, or operating environments, or the resulting data or information can be actively and accurately collected and used flexibly to make predictions.	S
	Installability	Degree of effectiveness and efficiency in which a product or system can be successfully installed and/or uninstalled in a specified environment (17-19)	R
	Replaceability	Degree to which a product can replace another specified software product for the same purpose in the same environment (17-19).	R
Explainability	Interactivity	Ability of the AI to provide effective and proactive interaction with the user (10,14,19).	R
	Transparency	Ability of the AI to provide effective and proactive interaction with the user (10,14,19).	R
	Verifiability	Ability of the AI and AIP, characterized by the degree of suitability for verification by various methods (10,14,19).	R
	Accountability	Ability of AIs to report in a defined form on the results of operations in a transparent manner (10,14,19).	R

Quality characteristics		Definition	CDW
Main	Sub		
	Interpretability	Ability of the AI to provide and interpret information in a user-friendly manner [10,14,19].	R
Trustworthiness	Diversity	Ability of the AI and AIPs to minimize failure risk to perform specified (defined as necessary) functions or tasks due to failures due to physical and informational factors, using various models, algorithms, and other means [10,14,19].	R
	Resiliency	Ability of the AI and AIP to continue to function amid changing requirements, parameters of the physical and information environment as well as the emergence of unspecified violations and failures. [10,14,19].	R
	Safety	Ability of the AI and AIP to avoid the risk of unacceptable damage and loss due to failures due to internal and external causes, and to reduce its consequences with the use of tools built in to the AI. [10,14,19].	R

(Table 7) Definition of the proposed ISO/IEC 25010 quality-in-use model encompassing AI

Quality Characteristics		Definition	CDW
Main	Sub		
Effectiveness	-	Degree of accuracy, and completeness in the achievement of planned results, with which users achieve specified goals [10,18,19]	R̂
Efficiency	-	Degree of accuracy and completeness with which users achieve goals considering the number of consumed resources [10,18]	R̂
Satisfaction	Usefulness	Degree to which a user is satisfied with their perceived achievement of pragmatic goals, including the results of use and consequences of use [10,18,19].	R
	Trust	Degree to which a user or other stakeholder has confidence that a product or system will function as intended [10,18,19].	R
	Pleasure	Degree to which a user obtains pleasure from fulfilling their personal needs [10,18,19].	R
	Comfort	Degree to which the user is satisfied with physical comfort [10,18,19].	R
	Specialization	Ability/degree of software to accurately respond to customer requests, actively reflect the latest information, and provide higher quality results than humans	S
	Personalization	Ability/degree to continuously monitor individual customers, analyze and accurately identify customer characteristics, and provide customized and optimized services to customers	S
Freedom form Risk	Economic Risk Mitigation	Degree to which a product or system mitigates the potential risk to financial status, efficient operation, commercial property, reputation, or other resources in the intended contexts of use [10,19,43].	R
	Health and Safety Risk Mitigation	Degree to which a product or system mitigates the potential risk to people in the intended contexts of use [10,18,19].	R
	Environmental Risk Mitigation	Degree to which a product or system mitigates the potential risk to property or the environment in the intended contexts of use [10,18,19].	R
	Security	(Including information and cyber security)– ability/ degree of software or platform to protect information and physical assets[14] so that other unidentified (unauthorized) persons or systems, including AI and AIPs, do not have access to them or have such access as specified type and level of authorization, including societal well-being, privacy, integrity, objectivity [14,19].	R̂

Quality Characteristics		Definition	CDW
Main	Sub		
Context Coverage	Context Completeness	Degree to which a product or system can be used with effectiveness, efficiency, freedom from risk, and satisfaction in all the specified contexts of use [10,18,19].	R
	Flexibility	Degree to which a product or system can be used with effectiveness, efficiency, freedom from risk, and satisfaction in contexts beyond those initially specified in the requirements [10,18,19].	R
Ethics	Fairness	Ability of the AI to minimize the risk of biased anomalies in ethical decisions (including lack of favoritism, discrimination on religious, racial, or other grounds, etc.), as well as misconceptions and errors in the modeling process [10,14,19].	R
	Graspability	Ability of the AI to provide the user with opportunities for the critical perception of AI in an open and democratic environment [10,14,19].	R
	Human agency	Ability of the AI to enable the user to make autonomous informed decisions about the use of AI [10,14,19].	R
	Human oversight	Ability of the AI to enable the user to control and, if necessary, interfere in a certain way with AI functioning [10.15,19].	R
	Redress	Ability of the AI to provide available [14] mechanisms to ensure adequate compensation for the effects of adverse effects on humans. [14][19].	R
Explainability	Causability	Ability of the AI to determine the cause-and-effect relationships between events that occur during its use [10,15,19].	R
	Completeness	Ability of the AI to be holistic in terms of compliance with all customer requirements [10,14,19].	R
	Comprehensibility	Ability of the AI to provide the user (or facilitate the user) with an understanding of the explanations sufficient to enable the use of the AI or the information obtained through it to perform other tasks [10,14,19].	R
	Interactivity	Ability of the AI to provide effective and proactive interaction with the user [10,14,19].	R
	Interpretability	Ability of the AI to provide and interpret information in a user-friendly way [10,14,19].	R
	Transparency	Ability of the AI to provide effective and proactive interaction with the user [10,14,19].	R
Trustworthiness	Acceptability	Ability of the AI to ensure at least partial compliance with customer requirements or consumer expectations [10,14,19].	R
	Accuracy	Ability of the AI and the AIP to ensure that the results of the requirements and/or functions presented by certain data are close to their true values [10,14,19].	R
	Resiliency	Ability of the AI and AI platform to continue to function amid changing requirements, parameters of the physical and information environment as well as the emergence of unspecified violations and failures [10,14,19].	R
	Robustness	Level/degree of functionality that allows the system to operate by maintaining performance levels over a wide range of input data and operating conditions, and reliably enter a shutdown state when data and conditions exceed specified limits	S
	Safety	Ability of the AI and AI platform to avoid the risk of unacceptable damage and loss due to failures due to internal and external causes, and to reduce its consequences with the use of tools built into the AI [10,14,19].	R
Lawfulness	-	Ability of the AI to comply with laws and regulations [10,14,19].	R
Responsibility	-	Ability of the AI to function considering the expectations of the client (user) by ethical norms, legal regulations, as well as to inform him in case of a possible violation [10,14,19].	R

4.4. Discussion and Implications

The ISO/IEC 25010 quality model that can encompass AI software, often referred to as the AI model, was developed based on the existing ISO/IEC standards related to AI and QCCC quality model research, which derived the quality standard items of the AI, AI platform, and AIS by compiling AI quality management systems [14]. Furthermore, the feasibility of the developed model was verified by applying a quality model derived from the quality metrics adopted in studies covering AI quality over the last five years. The study has the following implications:

First, among the conventional ISO/IEC 25010 quality models, the product quality model was expressed with 10 main quality attributes and 45 sub-quality attributes, whereas the quality-in-use model was expressed with 10 main quality attributes and 28 sub-quality attributes. This phenomenon indicates the necessity for considering the addition of the functional adaptability sub-quality attribute to the functional suitability main quality attribute, addition of user controllability and transparency to usability, and addition of robustness to robustness in the ISO/IEC 25010 product quality model. Furthermore, the security main quality attribute was reviewed as requiring changes. Specifically, the integrity sub-quality attribute revealed that the concept should be expanded to cover AI. Furthermore, although accountability is a sub-quality attribute existing in the ISO/IEC 25010 model, when the term is considered, the addition of intervenability, objectivity, and privacy sub-quality attributes should be considered. Particularly because accountability is included in the sub-quality attributes of the newly added explainability main quality attribute, a review of the description of the quality attribute revealed changing the accountability to traceability to be desirable. Moreover, two main quality attributes, namely explainability and trustworthiness, were added. First, the explainability main quality attribute requires the addition of interactivity, verifiability, accountability, and interpretability sub-quality attributes, and the trustworthiness main quality attribute requires the addition of diversity, resiliency, and safety subquantity attributes.

Next, considering the direction of revisions to the quality-in-use model, the freedom from risk main quality attribute requires the addition of security sub-quality attribute, which can include user-centered privacy, integrity, and objectivity. Furthermore, considering the addition of ethics, explainability, trustworthiness, lawfulness, and responsibility main quality attributes is necessary. Furthermore, ethics required the addition of fairness, graspability, human agency, human oversight, and redress sub-quality attributes, explainability required the addition of completeness, comprehensibility, interactivity, interpretability, and transparency, and trustworthiness required the addition of resiliency, robustness, safety, and accuracy. In particular, the validation process revealed that the satisfaction main quality attribute requires specialization and customization sub-quality attributes to measure satisfaction for AI expertise and satisfaction for personalization. Furthermore, the definition of some sub-quality attributes should be expanded so that the concept could encompass AI, even if the attribute name of the concept remained the same.

The second implication is that the quality-in-use of the ISO/IEC 25010 should be enhanced considerably. Thus, AI software, unlike ordinary software, has more quality control elements in terms of the user's authority, responsibility, and use of AI software. Thus, although the quality management of suppliers is important for conventional software, quality management capabilities of consumers is important for AI-based software.

Third, clarifying the elements that require expansion or change of the concept and defining the concept to determine in which direction the concept should be strengthened. For example, the concept of some sub-quality attributes in function suitability should be extended to include data. Although the ISO/IEC 25010 model is for software quality management and evaluation [49], AI software quality is determined by the learning and utilization of preprocessed training data through the process of source data, raw data, and labeling data. Data cannot be overlooked when assessing the functional accuracy or adequacy of ISO/IEC 25010 because product quality and quality-in-use depend on data. Therefore, redefining the concepts of sub-quality attributes by adding data items to the sub-quality attributes in functional suitability is desirable. The data quality attributes of ISO/IEC 25012 and 25024 [50] do not replace or deny the main quality attributes and sub-quality attributes defined in the model but should be defined and understood separately from them.

Fourth, the same quality attribute names should exist in the quality attributes required for both product quality and quality-in-use. In the conventional ISO/IEC 25010 quality model, the product quality attribute names and the quality-in-use attribute names were mutually exclusive. Thus, the quality attribute name alone made it clear whether it measured the product quality or the quality-in-use [51]. However, because the use of AI strengthened the quality control role on the user side, many requirements of product quality should be continuously valued and managed in quality-in-use. Therefore, explainability and trustworthiness were required in both quality models.

Fifth, quality items that require restructuring at the quality model level exist. For example, in the developed model, security is the main quality attribute in product quality, but in quality in use, it is dealt with as a sub-quality attribute of freedom form risk. Integrity, objectivity, and privacy are sub-quality attributes of security, but in terms of quality in use, they are all included in security. Furthermore, in product quality, robustness, the sub-quality attribute of reliability, is reduced to a sub-quality attribute of trustworthiness. Additionally, in the quality-in-use model, transparency is defined as a sub-quality attribute of the explainability main quality attribute. Accordingly, the concepts of quality attributes also should be defined differently according to the model.

This study investigated the quality items of an AI model and AI platform derived from the results of previous research. A method for deriving quality items requiring additional review was determined, and machine-learning-based quality items were established based on the goal being pursued by a project. However, the method for building such a model cannot be applied objectively. In this respect, it is believed that the research of [52,53] can provide insights into such applications. First, [52] provides insight into the further efforts that should be made for each development phase to improve quality when developing ML-based software from a software engineering perspective. Second, [53] presents a process for practitioners to ensure the quality of ML systems based on industrial cases. To apply the findings of this study to practically measure the quality of ML-based software, the research results of approaches described by Rahman et al. [52] and Siebert et al. [53] should also be considered.

5. Conclusions

In this study, an ISO/IEC 25010 quality model that encompasses AI software was developed and validated to address the limitations and problems arising from the application of ISO/IEC 25010, which is the most widely used international standard of software quality measurement and evaluation framework. The new ISO/IEC 25010 quality model was derived from the QCCC quality model in a previous study [14] that developed an AI-based quality model that can be applied to software, platforms, and systems. This study proposed the direction by validating and supplementing how the model should be modified to apply it comprehensively to AI software.

The study results suggested that the quality model should be reconstructed from the existing structure of 8 main quality attributes and 31 sub-quality attributes into 10 main quality attributes and 45 sub-quality attributes. In the case of a quality-in-use model, the existing structure of 5 main quality attributes and 9 sub-quality attributes needed to be reconstructed into 10 main quality attributes and 28 sub-quality attributes. Even if the attribute names are the same as existing standardized quality models, many attributes should be redefined to include AI software. Furthermore, reconstructing the AI to encompass the quality attributes established that the same quality attributes were required for the product quality and the quality in use. However, their hierarchy can be expressed differently in the positions of the main quality attributes and sub-quality attributes. Although some quality items are the same sub-quality attributes, they were mapped to different main quality attributes in the product quality model and quality-in-use model in some cases.

This study provides an excellent benchmarking example of what are the various aspects that models for software quality measurement and evaluation of the various international standards established to be applicable to general software should consider to evolve into a software quality model encompassing AI. The ISO/IEC 25010 application case revealed that the attributes and concepts of the quality model of the currently defined international standard should be revisited. The quality management and evaluation of AI software, which is rapidly spreading in many areas, is arbitrarily applied because unclear standards and requirements, which hinders the success of

AI-based projects. The quality model based on ISO/IEC 25010 developed [54] in this study can not only be used as the standard to solve these problems and also contribute significantly to accelerating the revision of existing international standards for software quality.

References

[1] S. Martínez-Fernández et al., "Software Engineering for AI-Based Systems: A Survey," ACM Transactions on Software Engineering and Methodology (TOSEM), Vol. 31, no. 2, pp. 1-59, Apr. 2022. https://doi.org/10.1145/3487043

[2] H. Jung, "The Software Quality Testing on the Basis of the International Standard ISO/IEC 25023," Journal of Korea Convergence Society, Vol. 7, no 6, pp. 35-41, Dec. 2016. https://doi.org/10.15207/JKCS.2016.7.6.035

[3] N. T. Nikolinakos, "EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies-The AI Act," 53. Springer, Jul. 2023. https://doi.org/10.1007/978-3-031-27953-9

[4] B. C. Stahl, D. Schroeder, & R. Rodrigues, "Ethics of Artificial Intelligence: Case Studies and Options for Addressing Ethical Challenges," Springer Nature, Switzerland, p.116, Cham. 2023. https://doi.org/10.1007/978-3-031-17040-9

[5] Johner Institut - Ihr Partner für Regularien im Bereich Medizinprodukte (johner-institut.de), https://www.johner-institut.de (accessed on 4 June 2023)

[6] ISO/IEC. Technical report: Software and systems Engineering－Software testing－Part 11: Guidelines on the testing of AI-based systems, ISO/IEC TR 29119-11:2020(E), First edition, Nov. 2020.

[7] ISO/IEC. Information Technology－Artificial Intelligence (AI)－Overview of Computational Approaches for AI Systems, 2021. Available online: https://www.iso.org/standard/78508.html (accessed on 4 June 2023). TR, 24372.

[8] ISO/IEC. Information Technology－Artificial Intelligence －Overview of Trustworthiness in Artificial Intelligence, 2020. Available online: https://www.iso.org/standard/77608.html (accessed on 4 June 2023). TR, 24028.

[9] ISO/IEC. Artificial Intelligence (AI). Assessment of the Robustness of Neural Networks. Overview. Available online: https://www.iso.org/standard/77609.html (accessed on 4 June 2023). TR, 24029-1:2021.

[10] UoN Digital Repository Home (uonbi.ac.ke), accessed on 4 Jul. 2023.

[11] M. Falco, E. Scott, & G. Robiolo, "Overview of an Automated Framework to Measure and Track the Quality Level of a Product," In 2020 IEEE Congreso Bienal de Argentina (ARGENCON), Resistencia, Argentina, pp. 1-7, Aug. 2021,. https://doi.org/10.1109/ARGENCON49523.2020.9505405

[12] S. B. Muradi, "Systematic Selection of Blockchain Platforms Using Fuzzy AHP-TOPSIS," University of Malaya (Malaysia) roQuest Dissertations Publishing, 30594302, Dec. 2022.

[13] M. Felderer, & R. Ramler, "Quality assurance for AI-based systems: Overview and challenges (introduction to inter-active session)," In Software Quality: Future Perspectives on Software Engineering Quality: 13th International Conference, SWQD 2021, Vienna, Austria, January 19-21, Jan. 2021, Proceedings 13 (pp. 33-42). Springer International Publishing. https://doi.org/10.1007/978-3-030-65854-0_3

[14] V. Kharchenko, H. Fesenko, & O. Illiashenko, "Quality Models for Artificial Intelligence Systems: Characteristic-Based Approach, Development and Application," Sensors, Vol. 22, no. 13, p. 4865, Jun. 2022. https://doi.org/10.3390/s22134865

[15] K.-H. Park, J. Kim, S. Lee, M.-K. Kim, K.-R. Lee, & J.-T. Han, "A Study on the Quality Control Method for Geotechnical Information Using AI," Journal of the Korean Geotechnical Society, Vol. 38, no. 11, pp. 87-95, Nov. 2022. https://doi.org/10.7843/kgs.2022.38.11.87

[16] ISO: Global standards for trusted goods and services. https://www.iso.org/home.html, accessed on June 2023.

[17] S.H. Kim, & W. J. Kim, "Analysis of Quality Correlation and Characteristic Importance of Application SW by SW Life Cycle Based on ISO 25000," Journal of KISS: Software and Applications, Vol. 41, no. 2, pp. 112-127, Apr. 2014.

[18] J. Estdale, & E. Georgiadou, "Applying the ISO/IEC 25010

Quality Models to Software Product," Systems, software and services process improvement, EuroSPI 2018, Bilbao, Spain, September 5-7, pp. 492-503, 2018. https://doi.org/10.1007/978-3-319-97925-0_42.

[19] ISO 25000 Software and Data Quality. https://iso25000.com (accessed on June 2023).

[20] M. Herrera, M. Á. Moraga, I. Caballero, & C. Calero, "Quality in use model for web portals (QiUWeP)," Current Trends in Web Engineering, ICWE 2010, Lecture Notes in Computer Science, vol 6385, Springer, Berlin, Heidelberg, 2010. https://doi.org/10.1007/978-3-642-16985-4_9

[21] UTokyo Repository. repository.dl.itc.u-tokyo.ac.jp, accessed on 4 Jul. 2023.

[22] S. Jayakumar, V. Sounderajah, P. Normahani, L. Harling, S. R. Markar, H. Ashrafian, & A. Darzi, "Quality Assessment Standards in Artificial Intelligence Diagnostic Accuracy Systematic Reviews: A Meta Research Study," NPJ Digital Medicine, Vol. 5, no. 1, p. 11, Jan. 2022. https://doi.org/10.1038/s41746-021-00544-y

[23] E. Manziuk, O. Barmak, I. Krak, O. Mazurets, & T. F. Skrypnyk, "Formal Model of Trustworthy Artificial Intelligence Based on Standardization," In IntelITSIS, Khmelnytskyi, Ukraine, pp. 190-197, Mar. 2021.

[24] T. Y. Kim, N. Ko, S. Yang, & S. M. Kim, "Trends in Network and AI Technologies," Electronics and Telecommunications Research Institute, Vol. 35, no. 5, pp. 1-13, Oct. 2020. https://doi.org/10.22648/ETRI.2020.J.350501

[25] G. Zhu, J. Zan, Y. Yang, & X. A. Qi, "Supervised Learning Based QoS Assurance Architecture for 5G Networks," IEEE Access, Vol. 7, pp. 43598-43606, Jul. 2019. https://doi.org/10.1109/ACCESS.2019.2907142.

[26] C. H. Baek, "A Study on Service Quality Evaluation in Artificial Intelligence Era Using Delphi Technique," Journal of Korea Service Management Society, Vol. 21, no. 3, pp. 1-15, Oct. 2020. https://doi.org/10.15706/jksms.2020.21.3.001

[27] D. Park, "Study on Service Quality Evaluation of AICC (Artificial Intelligence Contact Center) Considerations and Suggestions," JHSS, Vol. 13, no. 6, pp.4581-4592, Dec. 2022. https://doi.org/ 10.22143/HSS21.13.6.317.

[28] Q. Chen, Y. Gong, Y. Lu, & T. Tang, "Classifying and Measuring the Service Quality of AI Chatbot in Frontline Service," Journal of Bus. Res., Vol. 145, pp. 552-568, Jun. 2022. https://doi.org/10.1016/j.jbusres.2022.02.088

[29] C. H. Baek, S. U. Lim, & J. H. Choe, "A Study on Major Characteristic Analysis and Quality Evaluation Attributes of Artificial Intelligence Service," J Korean Soc Qual Manag, Vol. 47, no. 4, pp. 837-846, Dec. 2019. https://dx.doi.org/10.7469/JKSQM.2019.47.4.837

[30] C. H. Baek, S. Y. Kim, S. U. Lim, & J. Xiong, "Quality evaluation model of artificial intelligence service for startups," International Journal of Entrepreneurial Behavior & Research, Vol. 29, no. 4, pp. 913-940, May 2023. https://doi.org/10.1108/IJEBR-03-2021-0223

[31] S. Nativi, & N. S. DE, "AI Standardization Landscape: State of play and Link to the EC Proposal for an AI Regulatory Framework," Publications Office of the European Union, 2021. https://doi.org/10.2760/376602 (online)

[32] B. Y. Kim, "Trends and Implications of Regulatory Legislation in Major Countries as a Means of Controlling Artificial Intelligence," European Constitutional Law Research, vol 42, pp. 253-307, Aug. 2023. https://doi.org/10.21592/eucj.2023.42.253

[33] A. Idri, M. Bachiri, & J. L. Fernández-Alemán, "A Framework for Evaluating the Software Product Quality of Pregnancy Monitoring Mobile Personal Health Records," Journal of medical systems, Vol. 40, 50, pp. 1-17, Dec. 2016. https://doi.org/10.1007/s10916-015-0415-z

[34] M. Oriol, J. Marco, & X. Franch, "Quality models for web services: A systematic mapping. Inf. Software," Information and Software Technology, Vol. 56, no. 10, pp. 1167-1182, Oct. 2014. https://doi.org/10.1016/j.infsof.2014.03.012

[35] M. C. S. Fhang, & W. W. Tong, "Why a good process fail? Experience in building a sustainable and effective process for software development," ICSCA '18: Proceedings of the 2018 7th International Conference on Software and Computer Applications, pp. 40-45, Feb. 2018. https://doi.org/10.1145/3185089.3185107

[36] N. Condori-Fernandez, & P. Lago, "Characterizing the contribution of quality requirements to software

sustainability," Journal of systems and software, Vol. 137, pp 289-305, Mar. 2018.
https://doi.org/10.1016/j.jss.2017.12.005

[37] J. M. França, M. M. Costa Junior, & M. S. Soares, "Architecture-driven development of an electronic health record considering the SOAQM quality model," SN Computer Science, Vol. 1, no. 3, 140, Apr. 2020.
https://doi.org/10.1007/s42979-020-00150-x

[38] ISO/IEC 25059,
https://iso25000.com/index.php/en/iso-25000-standards/iso-25059/204-iso-iec-25059 (accessed on 1 July 2024).

[39] L. Sánchez-González, F. García, F. Ruiz, & M. Piattini, "Toward a quality framework for business process models," International Journal of Cooperative Information Systems, Vol. 22, no. 01, 1350003, 2013.
https://doi.org/10.1142/S0218843013500032

[40] C. S. Nam, "A Study on the Quality Monitoring and Prediction of OTT Traffic in ISP," The Journal of Korea Institute of Information, Electronics, and Communication Technology, Vol. 14, No. 2, pp.115-121, Apr. 2021.
https://doi.org/10.17661/jkiiect.2021.14.2.115

[41] H. G. Hwang, B. S. Kim, Y. T. Woo, Y. W. Yoon, S. C. Shin, & S. J. Oh, " A Development of Welding Information Management and Defect Inspection Platform Based on Artificial Intelligent for Shipbuilding and Maritime Industry," Journal of the Korea Institute of Information and Communication Engineering, Vol. 25, no. 2, pp. 193-201, Feb. 2021.
https://doi.org/10.6109/jkiice.2021.25.2.193

[42] D. M. Kim, "A study on the design of AI-based software quality testing framework using UML metadata," Department of IT Convergence (Artificial Intelligence Major) Graduate School, Dong-Eui University, 2022.

[43] H. S. Kim, "A Study on the Data Quality Management Evaluation Model," Journal of the Korea Convergence Society, Vol. 11, no. 7, pp. 217-222, Jul. 2020.
https://doi.org/10.15207/JKCS.2020.11.7.217.

[44] QA4AI Consortium. Guideline for the Quality Assurance of AI-Based Products. Available online:
http://www.qa4ai.jp/QA4AI. Guideline 2019, 201905.pdf (in Japanese).

[45] C. Yun, H. Shin, S. Y. Choo, & J. Kim, "An Evaluation Study on Artificial Intelligence Data Validation Methods and Open-Source Frameworks," Journal of Korea Multimedia Society, Vol. 24, no. 10, pp. 1403-1413, Oct. 2021.
https://doi.org/10.9717/kmms.2021.24.10.1403.

[46] C. H. Baek, "A Study on Major Service Items of Consumers and Companies Using Convergence Technology in the Intelligent Age," International Journal of Software Innovation (IJSI), Vol. 10, no. 1, pp.1-12. Jan. 2021.
https://doi.org/10.4018/IJSI.339884

[47] Ml. L. Berihun, (Ed.) Advances of Science and Technology: 9th EAI International Conference, ICAST 2021, Hybrid Event, Bahir Dar, Ethiopia, August 27-29, 2021. Proceedings, Part I (Vol. 411), Springer Nature.

[48] E. Ferko, "Towards a standards-based architecture for digital twins facilitating interoperability," Malardalen University (Sweden), 2023.

[49] T. L. Alves, P. Silva, & M. S. Dias, "Applying ISO/IEC 25010 Standard to prioritize and solve quality issues of automatic ETL processes," 2014 IEEE International Conference on Software Maintenance and Evolution, pp. 573-576, Sep. 2014.

[50] F. Gualo, M. Rodríguez, J. Verdugo, I. Caballero, & M. Piattini, "Data Quality Certification Using ISO/IEC 25012: Industrial Experiences," Journal of Systems and Software, Vol. 176, pp. 110938. Jun. 2021.
https://doi.org/10.1016/j.jss.2021.110938.

[51] D. Acton, "Assessing Quality in Software Engineering: A Pragmatic Approach," University of Pretoria (South Africa), 2013.

[52] M. S. Rahman, F. Khomh, A. Hamidi, J. Cheng, G. Antoniol, & H. Washizaki, "Machine learning application development: practitioners' insights," Software Quality Journal, Vol. 31, no. 4, pp. 1065-1119, Dec. 2023.
https://doi.org/10.1007/s11219-023-09621-9

[53] J. Siebert, L. Joeckel, J. Heidrich, A. Trendowicz, K. Nakamichi, K. Ohashi, & M. Aoyama, "Construction of a quality model for machine learning systems," Software Quality Journal, Vol. 30, no. 2, pp. 307-335, Jun. 2022.
https://doi.org/10.1007/s11219-021-09557-y

[54] D. Bán, "Static Source Code Analysis in Pattern Recognition," Performance Optimization and Software Maintainability (Doctoral dissertation, Szegedi Tudomanyegyetem (Hungary) 2018.

◐ 저 자 소 개 ◑

김 승 희(Seung-Hee Kim)
2003년 동국대학교 컴퓨터멀티미디어공학(공학사)
2005년 연세대학교 산업정보경영전공(공학석사)
2014년 서울과학기술대학교 산업정보시스템전공(공학박사)
2016년～2018년 한국기술교육대학교 컴퓨터공학부 교수
2018년～현재 한국기술교육대학교 IT융합소프트웨어공학과 교수
관심분야 : 소프트웨어 품질공학, AI기반 IT 프로젝트 관리, 블록체인, 일학습병행 등
E-mail : sh.kim@koreatech.ac.kr