

Original

Open Access

Prioritizing AI Standards: Integrating Requirements and Expert Criteria

Mehdi Azizi Mehmandoust ¹, Hamed Haji Molla Mirzaei ^{2*}

1. Ph.D. student in strategic management of cyber space, Supreme National Defense University and researcher at Innovation, ICT research Institute, Iran.

2. Faculty of Strategic Management of Cyberspace, Supreme National Defense University, Tehran, Iran.

Abstract

As artificial intelligence (AI) continues to advance across various sectors, the need for explicit and comprehensive standards governing its development, deployment, and use becomes increasingly critical. Establishing high-priority standards in the field of AI is essential to ensure responsible and ethical adoption of this transformative technology. While numerous international standards exist, prioritizing them and assessing their alignment with domestic requirements are crucial steps for effective utilization at the national level. To address this challenge, we propose a multi-dimensional approach for classifying AI standardization based on various criteria and the inherent priority of their respective domains. By considering stakeholder preferences, we can rank and prioritize AI standards effectively. Drawing inspiration from the European Union's method for prioritizing AI standards, we employ a systematic approach based on expert criteria and indicators extracted through text mining from existing standards. Our study identifies and introduces high-priority standards that serve as essential references for the development of national standards. These standards contribute to responsible AI implementation and foster ethical practices across diverse contexts.

Keywords Artificial Intelligence, Standardization, Text Mining, Prioritization Criteria, Expert Opinions.

1. Introduction

Artificial intelligence (AI) holds immense potential to revolutionize industries, foster innovation, and reshape our daily lives. With its wide-ranging applications across various domains, AI is poised to become an integral part of economic, social, and industrial sectors in the near future. Consequently, the identification, codification, and review of AI standards play a pivotal role in ensuring responsible development, deployment, and utilization of AI products [1, 2, 3].

The rapid evolution of AI technology, coupled with its dynamic nature, presents unique challenges that demand

continuous scrutiny and comprehensive regulatory frameworks [4, 5]. AI systems have the capacity to impact sensitive aspects of our communities, including privacy, security, fairness, transparency, accountability, and trustworthiness [6, 7]. Therefore, establishing a robust framework becomes imperative—one that mitigates potential risks, fosters trust, and safeguards the interests of diverse stakeholders [8, 9].

While numerous standards have been developed by private, national, and international standardization bodies [10], the task of identifying standards relevant to specific stakeholder needs necessitates prioritization strategies



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

* Correspondence:
ha.mirzaie@ut.ac.ir

within the vast landscape of existing standards. Prioritizing standards becomes a critical step when deploying global standards in domestic applications and formulating national AI standards. To achieve this, a comprehensive approach is essential—one that considers the perspectives of experts, policymakers, industry representatives, and end-users of AI systems [11].

By adopting such an inclusive approach, we can pave the way for responsible AI adoption and contribute to a harmonized global standardization landscape. When prioritizing AI standards, it is essential to consider a range of criteria that strike a delicate balance between fostering innovation and ensuring AI systems adhere to ethical principles and legal requirements. Additionally, internal priorities and preferences must be factored into the prioritization process [8, 12].

In this context, systematic indicators and step-by-step solutions play a pivotal role. These repeatable, measurable criteria are of paramount importance for prioritizing standards, as they enhance transparency, accuracy, and comprehensiveness. By minimizing errors and biases introduced by experts during the prioritization process, these indicators also facilitate automation [13].

Our research highlights a significant study conducted by the European Union (EU) on AI watch [14]. This study systematically prioritizes AI standards based on standard assessment indicators and examines the domains of related standards [14]. Drawing inspiration from [14], our report focuses on a set of international standards published by the International Organization for Standardization (ISO), the International Electrotechnical Commission (IEC), the Institute of Electrical and Electronics Engineers (IEEE), the European Telecommunications Standards Institute (ETSI), the British Standards Institution (BSI), and the International Telecommunication Union-Telecommunication Standardization Sector (ITU-T). Our work considers the specific requirements of domestic applications, encompassing trustworthiness, ethical considerations, responsibility, service quality assessment, and economic impact and public welfare.

The contribution of this paper is as follows: To achieve prioritization, we integrate information obtained through text mining indicators with qualitative criteria designed by experts. Ultimately, we identify nine high-priority standards from a pool of 28. These standards emphasize detailed design requirements and evaluation methods, covering aspects such as reliability, risk, privacy, security, safety, and bias.

In the subsequent section, we delve into the literature and elucidate the prioritization method employed in our study. The results obtained from this comprehensive approach contribute significantly to the field of AI standardization.

1.2. Literature Review

In the dynamic landscape of artificial intelligence (AI), establishing robust standards is essential. Systematic solutions and measurable indicators play a pivotal role in this process, ensuring transparency, accuracy, and comprehensive coverage. Additionally, automation streamlines the evaluation and prioritization of these standards. The AI Standards Hub website provides valuable insights into ongoing efforts [10]. Notably, two significant studies have tackled the automation challenge:

German Roadmap for AI Standardization [8]: This document outlines a strategic path for standardization in AI. It emphasizes systematic methodologies and criteria to guide the selection and prioritization of standards.

European Union's AI Watch Report [14]: The EU report proposes a comprehensive and rigorous approach to standardization. By integrating various stakeholders, it aims to harmonize AI practices across domains. This systematic framework ensures that standards align with societal needs, technological advancements, and ethical considerations.

In addition to these reports, several research studies delve into prioritizing AI use cases [15-17]. These investigations contribute to a holistic understanding of AI standardization, bridging the gap between theory and practical implementation [15-17].

In the following sections, we will delve deeper into the European Union's systematic approach and explore the nuances of prioritizing AI standards [14]. The German Roadmap for AI Standardization: The German Standardization Roadmap on Artificial Intelligence (AI), jointly published by DIN (Deutsches Institut für Normung) and DKE (Deutsche Kommission Elektrotechnik Elektronik Informations technik in DIN und VDE), serves as a critical guide in shaping AI standards [8]. Its key aspects are [8, 18]:

a) Importance of AI Standardization:

The roadmap underscores the significance of AI standardization. It advocates for a comprehensive review of existing documents and standards to assess their compatibility with AI applications. Ensuring compatibility involves identifying relevant standards and incorporating AI-related details into them [8].

b) Systematic Approach for Compatibility Assessment:

To achieve compatibility goals, a systematic approach is essential. Information technology tools and measurable criteria can be employed to examine various aspects of AI. The focus lies in assessing the relationship between existing documents and AI concepts [8].

c) SMART Standards: Toward Machine Interpretability:

The roadmap introduces the concept of SMART (Specif-

ic, Measurable, Achievable, Relevant, and Time-bound standards. These standards are transferable, understandable, and machine-readable. Other organizations, such as CEN/CENELEC and ISO/IEC endorse this approach as well. The ambitious goal is to establish machine-interpretable standards by 2030 [8].

d) Digital Transformation of Standards:

Currently, standards are mostly published in PDF format. However, the roadmap envisions a transformation: converting content into information that machines and applications can comprehend. This shift will enable the development of intermediary programs capable of accessing a readable repository of standards. The intermediary programs can automatically respond to user requests, either from individuals or from other programs. The intermediary programs act as experts in the field of the AI standards. They can partially compensate for the shortcomings of human expertise in AI standardization, especially in vertical domains. It would be possible to retrieve specific AI requirements or aspects with these systems in place [8].

e) Categorization of AI Standards:

The roadmap examines existing AI standards, ongoing projects, and standardization organizations. These entities are evaluated based on their association with seven AI topics [8]:

- Foundational principles
- Ethics and AI
- Quality
- Compliance measurement and certification
- IT security in AI systems
- Industrial automation
- Transportation and legislation
- AI in healthcare

Similar categorizations are also mentioned in [19]. Standards are initially divided into six main groups [19]:

- Foundational
- Platform/support
- Key technologies
- Products and services
- Applications
- Security/ethics

Each category is further dissected into subcategories [19].

f) Measurable Indicators and Future Directions:

While the roadmap provides valuable insights, a systematic approach based on measurable indicators has not been explicitly presented for these classifications [8]. In summary, the German Roadmap for AI Standardization charts a strategic course toward machine-interpretable

standards, leveraging technology, systematic approaches, and interdisciplinary collaboration. By bridging the gap between human expertise and AI, we pave the way for a standardized AI landscape that benefits society at large [8]. The European Union's study on AI watch: The European Union's study on AI watch [14] is the only study that introduces a methodical framework for assessing AI standards. This approach systematically incorporates indicators to measure standards, ensuring their relevance and effectiveness. Here are the key steps outlined in this study [14]:

a) General Study and Collection of Standards:

Initially, a comprehensive study of existing standards related to AI is conducted. Standards pertinent to AI are meticulously collected and reviewed [14].

b) High-Level Mapping to Desired Requirements:

In the second step, a high-level analysis takes place. Identified standards are mapped to specific requirements, such as [14]:

- Data and data governance
- Technical documentation
- Record-keeping
- Transparency and information for users
- Human oversight
- Accuracy
- Robustness and cybersecurity
- Risk management system
- Quality management system

This mapping relies on expert opinions and abstracts of the standards [14].

c) In-Depth Analysis and Prioritization:

The third step involves an in-depth analysis of standards. Numerical indicators are derived from direct textual alignment with the requirements. These indicators form the operationalization index. Additionally, other indicators, such as generality of domain, compliance with regulations, standard type, and maturity level, contribute to the suitability index. By systematically prioritizing standards using these indices, gaps can be identified [14].

d) Identifying Gaps and Providing Recommendations:

The fourth step focuses on potential gaps. Recommendations are offered to enhance existing standards. Overall, this systematic process ensures that AI standards align with evolving needs and technological advancements [14].

In summary, the European Union's approach emphasizes both rigor and practicality, enabling stakeholders to prioritize standards effectively and drive AI standardization forward [14].

Other relevant studies: In the realm of artificial intelli-

gence (AI), prioritizing use cases is crucial for efficient resource allocation and strategic decision-making. This subsection outlines systematic methods for identifying and ranking AI use cases, considering factors such as economic value, strategic alignment, and ease of implementation [15-17].

a) Clustering Approach:

The authors of [15] employ an iterative clustering approach to group closely related AI use cases. By analyzing factors such as required data, necessary AI capabilities, and applicable products or applications, this method identifies clusters with higher indicator values. These categories are evaluated based on value (economic value and strategic alignment), and ease of implementation (data, algorithm, process/system, and required expertise). Clusters with higher indicator values receive higher priority. Notably, this approach can be adapted to prioritize standardization efforts within the AI domain. By modifying specific criteria, it becomes a valuable tool for identifying standardization use cases.

b) Healthcare Domain Insights:

A comprehensive review study conducted in the healthcare domain provides valuable insights for successful AI application in healthcare organizations [16]. Key priorities include [16]:

- Data Governance: Establishing robust practices to ensure data quality, privacy, and security.
- Investment in Infrastructure: Allocating resources for novel data and analytical infrastructure.
- Skill Gap Mitigation: Addressing workforce skill gaps through training and upskilling.
- Strategic Collaborations: Fostering partnerships to enhance AI adoption.
- Transparency and Reproducibility: Leveraging good AI practices.
- Fairness and Bias: Explicitly addressing fairness concerns in AI applications.

These priorities align with best practices across various AI domains, emphasizing a holistic approach [16].

c) Measurable Criteria for Prioritization:

In another study, high-priority application areas for implementing AI technology were identified using the COPRAS method [17]. A panel of 150 experts collaborated to identify priority areas within the AI landscape. The following domains emerged as high-priority application areas [17]:

- Economic Applications: Leveraging AI for economic growth, productivity enhancement, and business competitiveness.
- Transportation: Enhancing transportation systems

through AI-driven solutions.

- Urban and Interurban Management: Optimizing urban infrastructure, traffic management, and city planning.
- Governance: Applying AI to improve government services, policy-making, and administrative efficiency.
- COPRAS (Comprehensive Proportional Assessment) employs multiple measurable criteria, including:

- Enhancing National Competitive Advantage: Aligning AI initiatives with national goals.
- Product Quality and Performance: Ensuring adherence to quality standards.
- Cost Reduction: Identifying cost-effective AI solutions.
- Human Resource Accessibility: Assessing ease of access to skilled personnel.
- Data Privacy Preservation: Prioritizing privacy-conscious practices.
- Energy and Water Consumption Optimization: Addressing sustainability.
- Fundamental and Applied AI Research: Encouraging innovation.
- Diversity of AI Applications: Fostering a wide range of use cases.
- Technical Knowledge Accessibility: Facilitating access to native expertise.
- Time Urgency: Considering time-sensitive applications.
- Market Demand: Responding to industry needs.
- Software and Hardware Infrastructure Accessibility: Ensuring necessary resources.
- Problem Detection and Corrective Actions: Building robust AI systems.
- Data Accessibility: Availability of relevant data.
- Industry-Academia Collaboration: Promoting knowledge exchange.
- Compatibility with Existing Technology Transfer Infrastructures: Seamless integration.
- Impact on Unemployment: Balancing workforce implications.
- Organizational Readiness: Preparing institutions for AI adoption.

By combining systematic methods, expert insights, and measurable criteria, organizations can holistically prioritize AI use cases. Strategic allocation of resources and alignment with broader goals will drive successful AI standardization efforts [17].

2. Material and methods

In this paper, we adopt a prioritization method informed by a systematic study conducted by the European Union [14]. Our goal is to prioritize standards based on measurable criteria, ensuring that each standard's overall value is

quantified relative to others. To achieve this, we employ two distinct categories of criteria:

a) Expert Criteria:

- Derived from input provided by domain experts, this category captures specific criteria relevant to standardization.
- Experts assign values to these criteria based on well-defined definitions. Their insights contribute to the overall prioritization process.

b) Text Mining Criteria:

- The second category emerges from text mining the documents associated with target standards.
- By analyzing these documents, we extract valuable information related to the standards' significance, applicability, and impact.

In this section, we delve into the details of each criterion, shedding light on how they collectively inform our systematic approach to prioritization.

2.1. Expert criteria

Eight criteria were considered by the experts in the field. The results of the evaluation of the experts are presented in Table 3.

Reliability (R): According to [20], reliability refers to a qualitative feature or ability of a system related to one of its critical performance metrics (positive or negative), such as accuracy, inaccuracy, accessibility, failure rate, responsiveness, etc. By adhering to the principles of reliability, we can ensure that the machine learning system functions as expected during deployment, avoids failure, and complies with the relevant execution constraints [21]. Sub-criteria are introduced in this study to further examine reliability. The sub-criteria include privacy, data governance and integrity, security, independence and control, diversity, non-discrimination and fairness, accountability, explainability and transparency, resilience, safety, stability and sustainability, objectivity (impartiality, scientificness), accuracy, auditability, representativeness, technical robustness and safety. **Ethics and Responsibility (EA):** Considering the impact of artificial intelligence applications on human users, adherence to ethics and responsibility is especially important in sensitive applications such as medicine, law, and finance [22-24]. Anthropocentricity, human agency and supervision, operationalization of ethical values, ethical design of privacy, and the consideration of the responsibilities of various stakeholders in the AI product market determine the presence of this criterion in an AI system. Additionally, monitoring mechanisms, accountability, and auditing and damage compensation ability provide tools to ensure compliance with

this principle.

Quality and Conformance Assessment (QCA): Assessing the quality and compliance of artificial intelligence systems with the expected performance criteria and existing regulations and requirements is of great importance [25]. The use of measurable criteria in the evaluation process helps improve reproducibility, mitigate bias, and control the system's performance, especially in the deployment phase. Some factors involved in estimating this criterion include: the development of successive testing and validation methods, testing requirements, models, and data, precise definition of evaluation criteria, continuous assessment of the system to understand critical cases, scalability, establishing a reference model for evaluation and interoperability, and designing practical, qualitative, and terminological standards.

Prioritized Application Domain (PAD): The closer the standard domain is to the specified priorities in laws, regulations, and policies, and the closer it is to domains with competitive advantages in the country, the higher the priority it has. Additionally, the domains of activity of the most important and largest AI companies in the country, as well as the areas with higher potential for adoption in the industry, have higher priorities in terms of practicality.

Economic Influence and Public Welfare (EPW): Considering the potential of artificial intelligence to impact various economic aspects, standards focusing on economically impactful domains and public services can be of higher importance. For example, standards related to the Internet of Things (IoT), industrial and financial standards, or standards related to the support of AI products can help improve service quality, facilitate economic processes, and ensure public welfare.

Domain and Application Generality (DAG): The broader the standard at the application level or domain level, (essentially, if a standard is horizontal), the more likely it covers broader concerns and can be more useful in more applications. Another sub-criterion that can be considered in the qualitative measurement of this index is the lack of standards related to the domain. The greater the gap between the requirements and current standards, the higher the priority of developing the standard.

Supply Chain Score (SCS): The supply chain includes data suppliers, computational and storage infrastructures, algorithm and model implementers, and software developers and deployers. They are directly involved in implementing standards and have the responsibility of com

pliance with standards in order to meet the requirements of all stakeholders. Therefore, it is necessary to measure how they score the quality and possibility of compliance with a standard and consider this in the prioritization of standards. Publisher Score (PS): This criterion determines the degree of credibility of a standard's verifier and publisher organization. Documents published by reputable organizations such as the ISO receive a higher score.

2.2. Text Mining criteria

In this step, a systematic method similar to the method proposed by the European Union [14] is used. This method utilizes text mining based on the keywords related to the requirements of this study to increase the accuracy of the analysis. For this purpose, an operationalization index (OI) is defined to match questions and documents, which is based on text mining methods. In the next step, the combination of the operationalization index with the qualitative criteria provided by the domain experts is used to create an overall suitability index. Higher priority standards can be identified based on this suitability index and other sub-criteria such as the overall operationalization or the operationalization of the requirements. In this study, based on expert opinions, four requirements, namely "reliability", "ethical considerations and responsibility", "service quality measurement", and "economic impact and public welfare", have been selected as important primary requirements. These requirements are broken down into finer sub-requirements listed in Table 1.

In this study, the operationalization index for each standard is defined based on a search for the sub-requirements in the text of the standard. Essentially, a set of sub-requirements is used to describe the requirement. To determine the relevance of each sub-requirement to the target standard, the TF-IDF score [26] between the standard text and the sub-requirement (as a question in retrieving information from documents) has been used. This score is calculated for each word based on its frequency in a specific document and its rarity in the entire corpus of the documents.

$$OI_k(s) = \frac{\sum_{l=1}^{n_k} \theta_{l_k}(s)}{n_k} \quad (2)$$

$$1 \leq l_k \leq n_k$$

If the TF-IDF score of a specific standard s for a sub-requirement l_k , (i.e. $TF-IDF(subreq.\#l_k, s)$) in Equation 3) is higher than the average score for all documents (i.e., $average(TF-IDF(subreq.\#l_k, s))$ in Equation 3), we

consider it relevant ($\theta_{l_k}(s) = 1$), otherwise, we consider it irrelevant ($\theta_{l_k}(s) = 0$). The rest of the calculations follow the Equations (1-3). The rest of the calculations follow the formulas (1-3). In these formulas, s refers to a standard in the target set of the study. k is an index corresponding to one of the four requirements mentioned in the previous paragraph, and $l(k)$ is the index of the sub-requirement among the n_k sub-requirements of requirement k .

$$OI(s) = \frac{\sum_{k=1}^4 OI_k(s)}{counter(s)} \quad (1)$$

$$counter(s) = |OI_k(s) \neq 0| \quad (2)$$

The goal of the operationalization index is to provide a radar chart for each requirement, showing which one of the analyzed standards is closer to the various goals of that requirement. Textual compatibility plays a role in the formula. Therefore, operationalization is defined as the existence of a common string of text between the standard and the description of the requirement. As a result, this method is compatible with the proposed solution for converting standards into machine-readable texts [8], as it utilizes text processing and information retrieval. It is important to note that in Equation (1), the operationalization index for the target standard is normalized by the number of related requirements (with non-zero operationalization), not by the total number of requirements. This choice ensures that the weights of standards that are only relevant to a limited subset of requirements are not reduced by uniform normalization across all standards. As a result, such standards are not unfairly penalized.

2.3. The Final criterion

In the end, the expert criteria and the text mining criterion are combined using a linear relationship to obtain the final criterion. A suitability index (SI) can be calculated using equation (4) for each standard. In this equation, w_n values are positive weight coefficients that adjust the importance of an aspect compared to the other aspects. The nine criteria of Reliability (R), Ethics and Responsibility (EA), Quality and Conformance Assessment (QCA), Prioritized Application Domain (PAD), Economic Influence and Public Welfare (EPW), Domain and Application Generality (DAG), Supply Chain Score (SCS), Publisher Score (PS) used in equation (4), are explained in details in section 2-1, By choosing the weight coefficients based on the perspectives of important stakeholders, different

$$\theta_{l_k}(s) = \begin{cases} 1 & TF-IDF(subreq.\#l_k, s) \geq average(TF-IDF(subreq.\#l_k, s)) \\ 0 & o.w. \end{cases} \quad (3)$$

Table 1. The considered criteria for systematic prioritization of standards

Sub-requirement	requirement
privacy, data governance, data integrity, security, autonomy, control, diversity, non-discrimination, justice, responsibility, explainability, transparency, resilience, stability, safety, objectivity, impartiality, neutrality, scientificness, avoidance of bias, accuracy, revisability, representative, robustness, reliability, repeatability, cybersecurity, regulatory compliance, accessibility	Reliability
human-centered, human oversight, value system, ethical design, accountability and governance, auditability, tradeoffs and redress	Ethical considerations and responsibility
technical requirements, quality backward chain, re-evaluation design, test criteria, risk management, risk assessment, critical state, interoperability, scalability, compatibility	Quality of service assessment
manufacturing, agriculture, transportation, health, education, trade, logistics, finance, home, government, public, security, city, environmental protection, court, big data, connected devices and Internet of Things, edge computing, smart sensors and data storage and transmission devices, natural language processing, intelligent speech, computer vision, recognition of biometrics features, virtual reality and augmented reality and human-computer interaction, transportation systems and vehicles, smart terminals and smart devices.	Economic Influence and Public Welfare

$$SI = \frac{w_1 \times OI + w_2 \times R + w_3 \times EA + w_4 \times QCA + w_5 \times PAD + w_6 \times EPW + w_7 \times DAG + w_8 \times SCS + w_9 \times PS}{\sum_{n=1}^9 w_n} \quad (4)$$

interpretations of the importance of the standards can be derived. Two examples of these interpretations and their results from the perspective of “AI product developers” and “regulators” are reported in Section 3.4.

3. Results

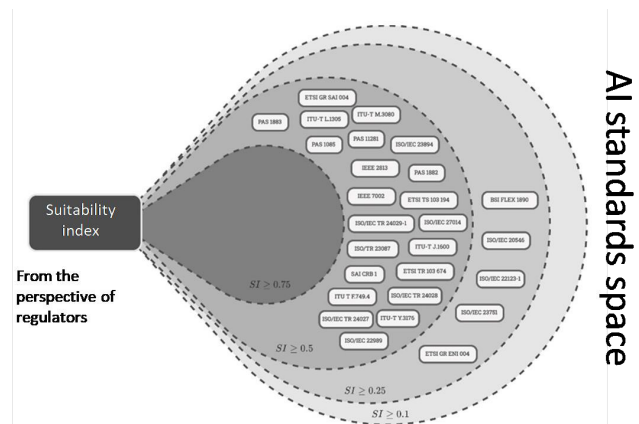
In this section, we first introduce the corpus of target standards. Then, we report the values of the expert criteria and the results obtained for the operationalization index. Finally, we calculate the suitability index using two different weighting methods. The results are presented in the form of tables and radar charts.

3.1. The population of studied standards

The population of initial standards used for prioritization in this study includes 28 standards, as shown in Table 2. The standards are selected based on the availability of their texts for calculating text mining-based criteria (refer to [10] for examples of the standards).

3.2. Prioritization considering the criteria of experts and operationalization

For each standard under study, the criteria introduced in section 3 are assigned values by five expert individuals in the field of artificial intelligence. These experts have PhD or master’s degrees in the fields of artificial intelligence, electrical engineering, or industrial engineering, and have

**Fig 1.** Radar diagram of operationalization index for the standards of the target population

more than three years of experience in the field of artificial intelligence. During this process, key information about the standards, including titles, general descriptions, objectives, scope of application, stakeholders, requirements, publishing process, and their operationalization are provided to the experts. Furthermore, the operationalization index is calculated through text mining using the equations in section (3-2) and is presented in Table 4. The radar chart of the standards for the overall operationalization index is shown in Figure 1. To categorize the standards into different priority levels, the following intervals are used: very

Table 2. Population of target standards in this study

Standard name	Standard ID	Standard name	Standard ID
Connected and automated vehicles – Vocabulary v2.0	BSI FLEX 1890	Information technology — Cloud computing and distributed platforms — Data sharing agreement (DSA) framework	ISO/IEC 23751
Experiential Networked Intelligence (ENI); Terminology for Main Concepts in ENI	ETSI GR ENI 004	Information technology — Artificial intelligence — Guidance on risk management	ISO/IEC 23894
Securing Artificial Intelligence (SAI); Problem Statement	ETSI GR SAI 004	Information security, cybersecurity and privacy protection — Governance of information security	ISO/IEC 27014
SmartM2M; Artificial Intelligence and the oneM2M architecture	ETSI TR 103 674	Information technology — Artificial intelligence (AI) — Bias in AI systems and AI aided decision making	ISO/IEC TR 24027
Network Technologies (NTECH); Autonomic network engineering for the self-managing Future Internet (AFI); Scenarios, Use Cases and Requirements for Autonomic/Self-Managing Future Internet		Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence	ISO/IEC TR 24028
		Artificial Intelligence (AI) — Assessment of the robustness of neural networks — Part 1: Overview	ISO/IEC TR 24029-1
		Automation systems and integration — The Big Picture of standards	ISO/TR 23087
		Use cases and requirements for multimedia communication enabled vehicle systems using artificial intelligence	ITU-T F.749.4
		Premium cable network platform – Framework	ITU-T J.1600
		Data centre infrastructure management system based on big data and artificial intelligence technology	ITU-T L.1305
		Framework of artificial intelligence enhanced telecom operation and management (AITOM)	ITU-T M.3080
		Machine learning marketplace integration in future networks including IMT-2020	ITU-T Y.3176
		Manufacturing – Establishing and implementing a security-minded approach – Specification	PAS 1085
		Connected automotive ecosystems – Impact of security on safety – Code of practice	PAS 11281
IEEE Standard for Big Data Business Security Risk Assessment	IEEE 2813	Data collection and management for automated vehicle trials for the purpose of incident investigation – Specification	PAS 1882
IEEE Standard for Data Privacy Process	IEEE 7002	Operational Design Domain (ODD) taxonomy for an automated driving system (ADS) – Specification	PAS 1883
Information technology — Big data — Overview and vocabulary	ISO/IEC 20546	Managing the Development of Artificial - Intelligence Software	SAE CRB 1
Information technology — Cloud computing — Part 1: Vocabulary	ISO/IEC 22123-1		
Information technology — Artificial intelligence — Artificial intelligence concepts and terminology	ISO/IEC 22989		

high priority for $0.7 \leq OI$, high priority for $0.5 \leq OI < 0.7$, medium priority for $0.3 \leq OI < 0.5$, low priority for $0.1 \leq OI < 0.3$, and very low priority for $OI < 0.1$. Two standards, ISO/IEC TR 24028 and ISO/IEC 22989, addressing trustworthiness in AI and AI concepts and terminology, respectively, have high operationalization, while the remaining standards have medium, low, or very low operationalization. As shown in Table 4, for each of the four requirements, “reliability” and “economic influence and public welfare” have two standards with high or very high priority, while “ethics and responsibility” has six standards with high or very high priority. For the last requirement, “service quality measurement,” seven prioritized standards are identified, although five of them have an operationalization value of 0.5, which is considered a threshold value. Therefore,

based on the standard corpus studied in this work, the two requirements of “trustworthiness” and “economic impact and public welfare” require more attention and the publication of more standards.

3.4. Final prioritization of standard based on different perspectives

In this section, the standards are prioritized based on the final criterion, which is the suitability index. The suitability index is a type integrating dataset resulting from expert opinions and text mining, and takes advantage of both types of criteria while mitigating their disadvantages by combining them. Experts can use more complex reasoning compared to text mining, but they might overlook some parts of the standard or existing features due to low accuracy and

Table 3. The Value of expert criteria for each standard (Values are assigned from 1 to 5, where 1 is minimum and 5 is maximum value)

Publisher Score	Spply Chain Score	Generality and diversity of Domain and Application	Economic Influence and Public Welfare	Prioritized Application Domain	Quality and Conformance Assessment	Ethics and Responsibility	Reliability	<i>Standard ID</i>
4	3	3	4	2	2	2	2	BSI FLEX 1890
4	3	3	3	2	2	2	2	ETSI GR ENI 004
4	3	3	4	2	3	4	4	ETSI GR SAI 004
4	3	3	3	3	3	3	3	ETSI TR 103 674
4	3	3	3	3	3	3	3	ETSI TS 103 194
4	4	4	3	4	4	4	4	IEEE 2813
4	5	5	3	4	4	4	4	IEEE 7002
5	3	3	3	2	2	2	2	ISO/IEC 20546
5	3	3	3	2	2	2	2	ISO/IEC 23123-1
5	3	3	3	2	2	2	2	ISO/IEC 22989
5	3	3	3	3	2	2	2	ISO/IEC 23571
5	4	4	3	3	4	4	4	ISO/IEC 23894
5	4	4	2	3	4	4	4	ISO/IEC 27014
5	5	5	3	2	3	5	5	ISO/IEC TR 24027
5	5	4	2	2	3	5	5	ISO/IEC TR 24028
5	5	5	2	4	5	4	4	ISO/IEC TR 24029-1
5	3	3	4	4	2	2	2	ISO/TR 23087
4	5	5	4	4	3	3	3	ITU-T F.749.4
4	3	3	3	3	3	3	3	ITU-T J.1600
4	3	3	3	3	3	3	3	ITU-T L.1305
4	3	3	3	3	3	3	3	ITU-T M.3080
4	3	3	3	3	3	3	3	ITU-T Y.3176
4	4	4	5	2	4	4	5	PAS 1085
4	4	4	3	3	4	4	4	PAS 11281
4	5	5	3	4	4	4	4	PAS 1882
4	3	3	4	4	3	3	3	PAS 1883
4	4	4	2	3	4	4	4	SAE CRB 1

Table 4. Resulted values for the operationalization and suitability indexes

<i>OI for "economic influence and public welfare"</i>	<i>OI for "quality of service assessment"</i>	<i>OI for "responsibility"</i>	<i>OI for "reliability"</i>	<i>OI interval</i>	<i>OI</i>	<i>SI interval (perspective 2)</i>	<i>SI (perspective 3)</i>	<i>SI (perspective 2)</i>	<i>SI (perspective 1)</i>	<i>Standard ID</i>
0.20	0.10	0.57	0.17	low	0.26	high	0.54	medium	0.47	BSI FLEX 1890
0.8	0.10	00.0	0.24	low	0.14	high	0.52	medium	0.43	ETSI GR ENI 004
0.48	0.30	0.57	0.48	medium	0.46	high	0.64	high	0.65	ETSI GR SAI 004
0.36	0.20	00.0	0.17	low	0.24	high	0.59	high	0.54	ETSI TR 103 674
0.28	0.50	0.14	0.10	low	0.26	high	0.60	high	0.55	ETSI TS 103 194
0.48	0.20	0.29	0.31	medium	0.32	very high	0.75	high	0.69	IEEE 2813
0.24	0.60	0.43	0.48	medium	0.44	very high	0.84	high	0.73	IEEE 7002
0.32	0.10	00.0	0.24	low	0.22	high	0.54	medium	0.45	ISO/IEC 20546
0.32	0.20	0.29	0.14	low	0.24	high	0.54	medium	0.45	ISO/IEC 22123-1
0.52	40.0	0.57	0.52	high	0.50	high	0.56	high	0.50	ISO/IEC 22989
0.32	0.20	00.0	0.31	low	0.28	high	0.56	medium	0.48	ISO/IEC 23751
0.24	0.50	0.71	0.31	medium	0.44	high	0.76	high	0.70	ISO/IEC 23894
0.28	0.50	0.29	0.28	medium	0.34	high	0.74	high	0.65	ISO/IEC 27014
0.24	0.30	0.71	0.28	medium	0.38	very high	0.84	high	0.72	ISO/IEC TR 24027
0.60	0.50	0.86	0.72	high	0.67	very high	0.81	high	0.73	ISO/IEC TR 24028
0.36	0.40	0.43	0.34	medium	0.38	very high	0.86	high	0.73	ISO/IEC TR 24029-1
0.32	0.30	0.29	0.10	low	0.25	high	0.59	high	0.52	ISO/TR 23087
0.36	0.10	0.43	0.14	low	0.26	very high	0.80	high	0.65	ITU-T F.749.4
0.32	0.00	0.00	0.24	low	0.28	high	0.60	high	0.55	ITU-T J.1600
0.16	0.10	0.14	0.24	low	0.16	high	0.59	high	0.53	ITU-T L.1305
0.20	0.50	0.00	0.07	low	0.26	high	0.60	high	0.55	ITU-T M.3080
0.04	0.20	0.00	0.03	very low	0.09	high	0.58	high	0.52	ITU-T Y.3176
0.28	0.10	0.14	0.24	low	0.19	very high	0.76	high	0.71	PAS 1085
0.36	0.70	0.43	0.38	medium	0.47	high	74.0	high	0.69	PAS 11281
0.24	0.20	0.00	0.34	low	0.26	very high	83.0	high	0.70	PAS 1882
0.20	0.00	0.14	0.17	low	0.17	high	62.0	high	0.58	PAS 1883
0.08	0.40	0.43	0.10	low	0.25	high	0.71	high	0.63	SAE CRB 1

Table 5. The set of weights corresponding to the priorities of AI developers

parameter	W_1	W_2	W_3	W_4	W_5	W_6	W_7	W_8	W_9
value	3	4	4	4	4	4	8	12	4

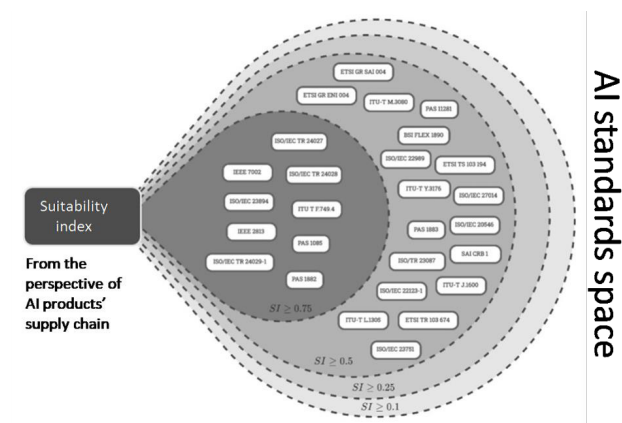


Fig 3. Radar diagram of the suitability index from the perspective of AI products' supply chain

given the most weight after the text mining criterion. The set of weights corresponding to these stakeholders is presented in Table 5.

Perspective 2: Artificial Intelligence Supply Chain

This group of stakeholders is directly involved in the implementation and execution of artificial intelligence standards. Their role is to create infrastructures, models, and AI-based applications while ensuring the fulfillment of the requirements of other stakeholders in the final system or the providing for their fulfillment in the post-production stage. The following values have been considered for the coefficients, where the supply chain score (SCS) has the highest weight compared to others. The set of weights corresponding to this group can be seen in Table 6.

Finally, based on these two weight categories, two types of suitability indexes are calculated and presented in Table 4. The prioritization ranges are similar to the operationalization index. For $0.75 \leq SI$, standards have very high priority, for $0.5 \leq SI < 0.75$, they have high priority, for $0.25 \leq CI < 0.5$, they have medium priority, for $0.1 \leq SI < 0.25$ they have low priority, and for $SI < 0.1$, they have very low priority. Radar diagrams in Figure 2 and

Figure 3 show the prioritization categories according to the perspectives. In general, nine standards were identified as high or very high priority standards by both perspectives:

- ISO/IEC TR 24027
- ISO/IEC TR 24028
- ISO/IEC TR 24029-1
- ISO/IEC 23894
- ITU-T F.749.4
- IEEE 7002
- IEEE 2813
- PAS 1085
- PAS 1882

4. Conclusion

In this study, artificial intelligence standards are evaluated from the perspective of domain experts and systematic text mining-based criteria. Following the European Union's study, we redefined the suitability and operationalization criteria for the standards and prioritized the corpus of available standards. Finally, considering the number of high and very high priority standards according to Table 4, among the desired requirements, "trustworthiness" and "economic influence and public welfare" had a larger gap with the target standards. This indicates the need for further publication of related standards or expansion of the target standard corpus to address these two requirements. Additionally, in terms of the operationalization index, ISO/IEC TR 24028 and ISO/IEC 22989, and in terms of suitability index, in common with perspectives of regulators and artificial intelligence supply chain, ISO/IEC TR 24027, ISO/IEC TR 24028, ISO/IEC TR 24029-1, ISO/IEC 23894, ITU-T F.749.4, IEEE 7002, IEEE 2813, PAS 1085, and PAS 1882, were identified as high priority standards. The content of these standards can be summarized as follows:

ISO/IEC TR 24027: This standard discusses bias in artificial intelligence systems and AI-based decision-making. It describes the measurement and evaluation methods for bias that are important in identifying weaknesses caused by bias in a system. Various stages of the system's life cycle, including data collection, training, continuous learning, design, testing, evaluation, deployment, and utilization, are examined in terms of bias within the framework of this document.

ISO/IEC TR 24028: This standard studies reliability in artificial intelligence systems. In other words, it explains methods for ensuring reliability in AI systems through transparency, explainability, and controllability. It discusses engineering errors and common risks associated

with AI systems and provides solutions to mitigate them. It also addresses methods of evaluation and achieving accessibility, flexibility, trustworthiness, accuracy, safety, security, and privacy in AI systems.

ISO/IEC TR 24029-1: This standard examines various existing methods for assessing risks related to the reliability of neural networks. These methods are widely used in industry, governance, and education. The standard provides the necessary background to understand these methods.

ISO/IEC 23894: This standard discusses risk management in the development, production, deployment, and utilization of AI services and systems by organizations. It explains effective implementation strategies and the application of risk management in AI-related activities within a system, which can be customized to fulfill the organization's needs.

ITU-T F.749.4: This standard addresses the applications and requirements of AI-based transportation systems equipped with multimedia communication. It includes general aspects, high-level architecture, network and service requirements, as well as functional and non-functional requirements.

IEEE 7002: This standard includes the requirements of an engineering process for applying privacy considerations to products, services, and systems involved with the personal data of employees, customers, and external users that flow throughout the system's life cycle. By relying on a use case and model, this standard introduces procedures, diagrams, and a list of tasks that allow users assessing product compliance with privacy guidelines.

IEEE 2813: This standard describes methods for assessing security risks in user behavior, applied analysis layer, and fundamental analysis layer for big data.

PAS 1085: This standard considers the requirements for security management in construction and manufacturing organizations. Its goal is to protect organizational liability and reputation, intellectual property, safety and security of production privileges, and the integrity of manufactured products. In addition, some methods for meeting the requirements in the standard are examined.

PAS 1882: This standard focuses on the requirements for collecting, maintaining, storing, and sharing data during tests of autonomous vehicles.

The selected standards are mostly concerned with spec-

ifying requirements and evaluation methods for issues including reliability, system risk, privacy, security, safety, and bias. Detailed requirements and the existence of measurement criteria can help assess the quality and compliance of systems with the standards. Using the measurement criteria is essential for adhering to the standards, since without quantifying a system's strengths and weaknesses in different domains, its compliance with relevant standards cannot be ensured. Moreover, these methods can help identify areas in which the system needs improvement and can reveal the current gaps between the existing system and the standard system. In the next step, by selecting and applying appropriate standards and enhancing AI systems, these gaps can be addressed.

The approach used in this paper is presented merely as an example of using this systematic solution for prioritizing standards and introducing its domestic cases. It is necessary to repeat this study with a larger population of standards and focusing on the requirements resulting from a conversation with stakeholders to achieve a more precise refinement for national purposes.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: Jan. 2024 Accepted: Jun. 2024

Published online: Jul. 2024

DOI: 10.22034/ASAS.2024.440428.1050

References

- [1] E. Blasch, J. Sung, T. Nguyen, C. Daniel, and A. Mason, "Artificial Intelligence Strategies for National Security and Safety Standards," 11/07 2019.
- [2] Laux, Johann and Wachter, Sandra and Mittelstadt, Brent, Three Pathways for Standardisation and Ethical Disclosure by Default under the European Union Artificial Intelligence Act (February 20, 2023). Available at SSRN: <https://ssrn.com/abstract=4365079> or <http://dx.doi.org/10.2139/ssrn.4365079>
- [3] Zielke, T. (2020). Is artificial intelligence ready for standardization?. In Systems, Software and Services Process Improvement: 27th European Conference, EuroSPI 2020, Düsseldorf, Germany, September 9–11, 2020, Proceedings 27 (pp. 259–274). Springer International Publishing.
- [4] Roberts, H., Cows, J., Morley, J. et al. The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation. *AI & Soc* 36, 59–77 (2021). <https://doi.org/10.1007/s00146-020-00992-2>
- [5] Ebers, Martin, Standardizing AI - The Case of the European Commission's Proposal for an Artificial Intelligence Act (August 6, 2021). The Cambridge Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics, Available at SSRN: <https://ssrn.com/abstract=3900378> or <http://dx.doi.org/10.2139/ssrn.3900378>
- [6] A. HLEG, "ETHICS GUIDELINES FOR TRUSTWORTHY AI," High-Level Expert Group on AI 2019.
- [7] Um, T.-W., Kim, J., Lim, S., & Lee, G. M. (2022). Trust Management for Artificial Intelligence: A Standardization Perspective. *Applied Sciences*, 12(12), 6022. MDPI AG. Retrieved from <http://dx.doi.org/10.3390/app12126022>
- [8] W. Wahlster and C. Winterhalter, "German standardization roadmap on artificial intelligence," German Institute for Standardization and DKE German Commission for Electrical, Electronic & Information Technologies of DIN and VDE November, 2020
- [9] Ziegler, W. (2020). A Landscape Analysis of Standardisation in the Field of Artificial Intelligence. *Journal of ICT Standardization*, 8(2), 151–184. <https://doi.org/10.13052/jicts2245-800X.824>
- [10] AI Standards Hub, The Alan Turing Institute, accessed July 19, 2023 <<https://aistandardshub.org/>>
- [11] Miller, G. J. (2022). Stakeholder roles in artificial intelligence projects. *Project Leadership and Society*, 3, 100068
- [12] "Discerning signal from noise: The state of global AI standardization and what it means for Canada," Schwartz Reisman Institute for Technology and Society at the University of Toronto 2023.
- [13] Jenkins, G. M., & Youle, P. V. (1968). A systems approach to management. *Journal of the Operational Research Society*, 19(sup1), 5–21.
- [14] J. Soler Garrido et al, "AI Watch: Artificial Intelligence Standardisation Landscape Update," 2023, issue 1831–9424 (online)
- [15] Applying AI: How to find and prioritize AI use cases, Applied AI
- [16] Fisher, S., Rosella, L.C. Priorities for successful use of artificial intelligence by public health organizations: a literature review. *BMC Public Health* 22, 2146 (2022). <https://doi.org/10.1186/s12889-022-14422-z>
- [17] Safari, E., & Ebrahimi, K. (2022). Prioritizing application areas to deploy artificial intelligence technology using thematic analysis and COPRAS. *Production and Operations Management*, 13(4), 91–110. doi: 10.22108/pom.2023.134630.1457

- [18] BMWi, StrategieKünstlicheIntelligenz der Bundesregierung. Berlin, 2018. (June 25, 2023). Available at: www.ki-strategie-deutschland.de
- [19] Murphy, B., & Lead, C. T. (Eds.). (2020). Artificial Intelligence Security Standardization White Paper (2019 Edition).
- [20] Bosnić, Z., & Kononenko, I. (2009). An overview of advances in reliability estimation of individual predictions in machine learning. *Intelligent Data Analysis*, 13(2), 385-401.
- [21] Saria, S., & Subbaswamy, A. (2019). Tutorial: safe and reliable machine learning. *arXiv preprint arXiv:1904.07204*.
- [22] Lechterman, Theodore M. (Forthcoming). The concept of accountability in ai ethics and governance. In J. Bullock, Y.C. Chen, J. Himmelreich, V. Hudson, A. Korinek, M. Young, and B. Zhang (Eds.), *The Oxford handbook of AI governance*. Oxford University Press.
- [23] Unsworth, H., Wolfram, V., Dillon, B., Salmon, M., Greaves, F., Liu, X., ... & Weller, A. (2022). Building an evidence standards framework for artificial intelligence-enabled digital health technologies. *The Lancet Digital Health*, 4(4), e216-e217.
- [24] Baxter, S. L., & Lee, A. Y. (2021). Gaps in standards for integrating artificial intelligence technologies into ophthalmic practice. *Current Opinion in Ophthalmology*, 32(5), 431-438.
- [25] Siebert, J., Joeckel, L., Heidrich, J. et al. Construction of a quality model for machine learning systems. *Software Qual J* 30, 307–335 (2022). <https://doi.org/10.1007/s11219-021-09557-y>.
- [26] Manning, C. D. (2009). *An introduction to information retrieval*. Cambridge university press.
- [27] Mitchell, R. K., Agle, B. R., & Wood, D. J. (1997). Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *Academy of management review*, 22(4), 853-886.

Submit your manuscript to Advances in the standards and applied sciences journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open Access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

**Submit your next manuscript at:
journal.standards.ac.ir**