



ARTICLE

Computational Assessment of Information System Reliability Using Hybrid MCDM Models

Nurbek Sissenov^{1,*}, Gulden Ulyukova^{1,*}, Dina Satybaldina² and Nikolaj Goranin³

¹Department of Information Security, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

²Research Institute of Information Security and Cryptology, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

³Department of Information Systems, Vilnius Gediminas Technical University, Vilnius, Lithuania

*Corresponding Authors: Nurbek Sissenov. Email: nurbek0692@gmail.com; Gulden Ulyukova. Email: gulden1590@gmail.com

Received: 03 November 2025; Accepted: 12 January 2026; Published: 12 March 2026

ABSTRACT: The reliability of information systems (IS) is a key factor in the sustainable operation of modern digital services. However, existing assessment methods remain fragmented and are often limited to individual indicators or expert judgments. This paper proposes a hybrid methodology for a comprehensive assessment of IS reliability based on the integration of the international standard ISO/IEC 25010:2023, multicriteria analysis methods (ARAS, CoCoSo, and TOPSIS), and the XGBoost machine learning algorithm for missing data imputation. The structure of the ISO/IEC 25010 standard is used to formalize reliability criteria and subcriteria, while the AHP method allows for the calculation of their weighting coefficients based on expert assessments. The XGBoost algorithm ensures the correct filling of gaps in the source data, increasing the completeness and reliability of the subsequent assessment. The resulting weighted indicators are aggregated using three MCDM methods, after which an integral reliability indicator is formed as a percentage. The methodology was tested on six real-world information systems with different architectures. The results demonstrated high consistency between the ARAS, CoCoSo, and TOPSIS methods, as well as the stability of the final rating when the criterion weights vary by $\pm 10\%$. The proposed approach provides a reproducible, transparent, and objective assessment of information system reliability and can be used to identify system bottlenecks, make modernization decisions, and manage the quality of digital infrastructure.

KEYWORDS: Information system; reliability; ISO/IEC 25010:2023; multi-criteria method; ARAS; CoCoSo; TOPSIS; AHP; machine learning; extreme gradient boosting

1 Introduction

Modern information systems (IS) form the foundation of digital infrastructure in government, commercial, and mission-critical industries. The reliability of such systems directly impacts business continuity, data security, and user trust. Despite a significant amount of research devoted to assessing the quality and reliability of IS, existing approaches remain fragmented and focused either on individual technical indicators or subjective expert assessments [1–3].

Most modern studies use either standardized quality models without quantitative formalization or multi-criteria decision making (MCDM) methods without strict adherence to international standards [4]. Furthermore, a significant portion of the studies do not take into account the problem of incomplete data, which is typical for real-world IS, especially when analyzing operational and failure indicators [5,6].

This significantly reduces the reliability of the final assessment and limits the practical applicability of the proposed methods.

This study proposes a comprehensive approach aimed at overcoming these limitations by integrating the international standard ISO/IEC 25010:2023, hybrid MCDM models, and machine learning methods. This synthesis not only allows for the structuring of reliability as a multidimensional property but also provides a quantitative, reproducible, and robust assessment adapted to the real-world operating conditions of information systems.

1.1 ISO/IEC 25010:2023 as a Normative Basis for Assessing the Reliability of Information Systems

The ISO/IEC SQuaRE (Software Product Quality Requirements and Evaluation) series of international standards has been the basis for formalized quality assessment of software products and information systems for several decades. The evolution of this series of standards began with the ISO/IEC 9126 model, which first proposed a systematic representation of software quality as a set of characteristics and subcharacteristics [7]. Despite its widespread use, ISO/IEC 9126 had several limitations related to insufficient detail in its metrics and poor adaptation to complex, distributed information systems.

To address these shortcomings, ISO/IEC 9126 was replaced by ISO/IEC 25010, which became a key element of the SQuaRE family. It offered a more rigorous and logically structured quality model, focusing not only on software but also on information systems as a whole [8]. The latest edition of the standard, ISO/IEC 25010:2023, reflects the modern requirements of digital transformation, including issues of reliability, security, performance, and resilience of systems under high load and distributed architecture.

ISO/IEC 25010:2023 defines nine core quality characteristics, including suitability, performance, compatibility, usability, reliability, security, maintainability, and portability [8,9]. A key advantage of this model is its hierarchical structure, which allows abstract quality properties to be decomposed into subcharacteristics and then into measurable indicators. This makes the standard a universal basis for formalizing complex quality properties of information systems.

The ISO/IEC 25010:2023 standard was selected for this study because it is the most relevant and internationally recognized model for assessing the quality of information systems, widely used in both scientific research and industrial practice [4,10]. Unlike many applied models, ISO/IEC 25010 ensures conceptual integrity and consistency of criteria, which is especially important for a comprehensive assessment of the reliability of information systems. The ISO/IEC 25010:2023 standard itself does not offer a quantitative mechanism for an integrated assessment of reliability. It defines what should be assessed, but does not determine how to aggregate indicators into a single numerical indicator. In most existing studies, the standard is used only as a descriptive or auxiliary model without rigorous mathematical formalization [4,9]. This creates a significant methodological gap, especially for comparative analysis and management decision-making.

Within the framework of ISO/IEC 25010, reliability is considered a comprehensive property, including subcharacteristics such as availability, fault tolerance, maturity, and recoverability. However, in existing studies, these subcharacteristics are typically analyzed in isolation or at a qualitative level, without developing a comprehensive numerical indicator of information system reliability [2,3].

In this study, ISO/IEC 25010:2023 is used not only as a conceptual model but also as a formal basis for developing a computational methodology for reliability assessment. The standard's hierarchical structure is used to create a three-level model of "criterion, subcriterion, and measurable indicator" which allows for linking the standard's requirements to actual operational data of information systems. This approach ensures a rigorous interpretation of reliability in accordance with international standards and simultaneously creates a foundation for further quantitative analysis.

The use of ISO/IEC 25010:2023 in this work is justified and methodologically sound. It serves as a fundamental framework for integrating multicriteria analysis and machine learning methods, enabling a transition from declarative quality assessment to a reproducible and quantifiable assessment of information system reliability.

1.2 Multicriteria Decision Making Models (MCDM)

Multi-criteria decision making methods are widely used to analyze complex systems characterized by the presence of many heterogeneous and often conflicting criteria [11,12]. Classical MCDM approaches, such as AHP, TOPSIS, VIKOR, ELECTRE, and DEMATEL, have become widely used in problems of assessing the efficiency, risk, quality, and sustainability of various technical and organizational systems [13–16].

The evolution of MCDM methods in recent years has led to the active development of hybrid approaches aimed at increasing the stability of results, reducing the subjectivity of expert assessments, and taking into account the interrelations between criteria [17–19]. The most common combinations are AHP–TOPSIS, AHP–DEMATEL, FAHP–VIKOR, and DEMATEL–ANP, which allow for consideration of both the hierarchical structure of criteria and their mutual influence [19,20]. Such hybrid models have proven their effectiveness in tasks such as supplier selection, investment project evaluation, and risk analysis.

Hybridization of MCDM methods has become a trend over the past decade, aimed at overcoming the limitations of individual methods. A literature review identifies several key types of hybrid approaches and their applications in related fields:

1. Combinations for accounting for uncertainty: Integration of fuzzy set theory methods with MCDM. For example, F-AHP in combination with F-TOPSIS or F-VIKOR allows for the effective formalization of linguistic expert assessments. The study [15] successfully applied FAHP-Interval TOPSIS to IT strategy selection, increasing the solution’s robustness to subjectivity. Similarly, in [21], the use of AHP-FTOPSIS for bank evaluation allowed for the integration of qualitative and quantitative criteria;

2. Combinations for analyzing criteria interdependencies: Combining weighting methods (AHP) with structural modeling methods (DEMATEL, ANP). AHP-DEMATEL is used when the criteria are not independent. Such a combination helped not only to determine the weight of supplier evaluation criteria but also to identify key causal factors. A more complex DEMATEL-ANP-VIKOR chain was used to select Six Sigma projects, taking into account the mutual influence and feedback between requirements;

3. Combinations for decision aggregation: Using several ranking methods (e.g., TOPSIS, VIKOR, WASPAS) followed by synthesis of their results to obtain a more robust final ranking.

A literature review (Table 1) shows that existing hybrid methods successfully address the challenges of accounting for uncertainty or interrelationships among criteria. However, they rarely combine standardized quality models with machine learning capabilities. This paper aims to fill this gap.

Table 1: Detailed comparison of modern hybrid MCDM approaches.

Research	Combination of Methods	Scope of Application	The Key Problem to be Solved	Advantages of the Approach	Disadvantages and Limitations	Parameters/ Features
Hemmati et al. (2019) [15]	FAHP + Interval TOPSIS	Choosing an IT service strategy	Accounting for uncertainty and imprecision in expert judgments	Improved robustness to input fuzziness; interval numbers improve flexibility.	Does not address the issue of missing data; difficulty in interpreting interval results for end users.	Using interval arithmetic; agreement between experts is controlled by the IS.

(Continued)

Table 1 (continued)

Research	Combination of Methods	Scope of Application	The Key Problem to be Solved	Advantages of the Approach	Disadvantages and Limitations	Parameters/Features
Etraj & Jayaprakash (2018)	AHP + DEMATEL	Green Supplier Assessment	Taking into account cause-and-effect relationships and interdependencies of criteria	Allows you to identify key drivers and resulting factors, not just their weight.	Computational complexity; requires deep understanding of the system by experts to construct the influence matrix.	Construction of a cause-and-effect map; calculation of a threshold value to filter out insignificant relationships.
Ortiz et al. (2020)	DEMATEL + ANP + VIKOR	Selecting Six Sigma Projects in Healthcare	Comprehensive consideration of mutual influences and feedback between criteria of different levels	The most complete modeling of complex systems with a network structure of dependencies.	Very high computational and expert load; risk of model overfitting.	Using ANP supermatrices; a compromise decision based on group utility and individual regret (VIKOR).
Beheshtinia et al. (2020)	AHP + FTOPSIS, AHP + FVIKOR	Evaluation of bank performance	Integrated assessment based on a variety of financial and non-financial criteria	Comparison of results of two methods (FTOPSIS and FVIKOR) for validation; accounting for uncertainty.	No single standardized system of criteria is used (e.g., ISO); it depends on the completeness of the source data.	Application of triangular fuzzy numbers; sensitivity analysis to weight changes.
Gupta et al. (2025)	Fuzzy Delphi + AHP + CoCoSo	Supply chain management in the food industry	Identifying barriers and strategies for sustainable development	A combination of expert consensus (Delphi), weighting (AHP) and compromise ranking (CoCoSo).	Long iterative process; focus on qualitative barriers rather than quantitative system metrics.	Multi-stage methodology; using fuzzy numbers to process qualitative judgments

In the field of information systems, MCDM methods are used primarily for comparative analysis of alternatives, including the selection of software platforms, cloud services, and architectural solutions [13,22]. In a number of studies, MCDM is used to evaluate individual aspects of IS quality, such as performance, security, or usability [4,23]. However, in most cases, reliability is considered either as one of the secondary criteria or as an abstract property without a rigorous quantitative interpretation.

An analysis of existing research shows that MCDM methods are rarely used for direct quantitative assessment of information system reliability in accordance with international quality standards. In the vast majority of studies, the result of applying MCDM is the ranking of alternatives, rather than the calculation of an absolute numerical reliability indicator suitable for monitoring, comparison with standard values, or management decision-making [21,24].

Preliminary studies utilized the modern MOBAC and CODAS methods, which have attracted the attention of researchers in recent years due to their robustness to changes in criterion weights and computational simplicity. Despite their analytical advantages, the results of applying these methods showed that they are primarily focused on obtaining the relative preference order of alternatives. The resulting values generated

by MOBAC and CODAS lack direct interpretability and do not allow for representing the reliability of an information system as an unambiguous quantitative indicator.

This aspect is a significant limitation in assessing the reliability of information systems, as practical tasks require not only comparing systems with each other but also determining their compliance with specified requirements or benchmarks. The lack of an absolute assessment scale complicates the analysis of reliability dynamics over time and the development of recommendations for system upgrades.

Unlike the aforementioned approaches, this study examines MCDM methods not as a standalone ranking tool, but as a component of a hybrid computational framework. The ARAS, CoCoSo, and TOPSIS methods are used together to generate an integrated reliability indicator, normalized as a percentage. This approach allows for the integration of the advantages of various MCDM models, mitigating the impact of their individual limitations, and moving from relative ranking to a quantitative assessment of information system reliability.

An analysis of the evolution and application of MCDM approaches shows that, despite their widespread use, existing methods do not provide a comprehensive quantitative assessment of IS reliability. This confirms the need to develop a hybrid methodology integrating MCDM with a formalized quality model and additional computational tools, as implemented in this study.

1.3 Machine Learning as a Tool for Ensuring Data Integrity

Practical assessment of information system reliability inevitably faces the problem of incomplete source data. Performance indicators such as MTBF, MTTR, failure rate, and resource utilization often contain gaps due to monitoring limitations, logging failures, or organizational factors [5]. Using such incomplete data in multicriteria analysis leads to distortion of the final assessments and a decrease in their reliability, making the task of data recovery critical.

In recent years, machine learning methods have been actively used to improve the accuracy of reliability analysis and the processing of complex, nonlinear dependencies in technical and information systems [25,26]. Unlike traditional statistical imputation methods, machine learning allows for the consideration of hidden relationships between indicators, thereby generating more realistic and consistent input information for subsequent analysis [6]. In this study, machine learning is used specifically as a tool for ensuring data integrity, rather than as a replacement for analytical or expert methods.

Various approaches have been considered for missing value imputation, including linear regression, k-NN imputation, random forests, and ensemble methods. Linear models have low computational complexity but are unable to adequately model nonlinear relationships between reliability metrics. k-NN methods are sensitive to the choice of distance metric and data scaling, while random forests, despite their high accuracy, demonstrate lower robustness with limited sample size [27].

In this study, we selected the XGBoost algorithm, which belongs to the class of gradient boosting and combines high accuracy with robust results. Compared to alternative methods, XGBoost effectively handles tabular data, models nonlinear relationships, and contains built-in regularization mechanisms that reduce the risk of overfitting [28,29]. Several comparative studies show that XGBoost outperforms traditional imputation methods in terms of reconstruction accuracy and robustness when working with incomplete and heterogeneous data [30].

Using XGBoost as a data recovery tool is a methodologically sound choice. Its integration into the proposed hybrid model eliminates the key limitation of MCDM methods—the requirement for a complete decision matrix, and thus ensures a more accurate, reproducible, and objective assessment of information system reliability.

1.4 Comparative Analysis of Existing Methods and Software Tools for Assessing the Reliability of Information Systems

In addition to methodological approaches, an important aspect is the existing software tools and practical methods used to assess IS reliability. Their analysis allows us to understand the current technological landscape and identify gaps that the proposed hybrid methodology fills.

Existing solutions can be roughly divided into several categories:

Monitoring and APM (Application Performance Management) tools: These include platforms such as Datadog, Dynatrace, New Relic, Nagios, and Zabbix. These systems provide extensive real-time data on availability, response time, resource utilization (CPU, RAM), error rates, and transactions. They are effective for collecting metrics and detecting incidents, but their focus is biased toward performance and observability rather than a comprehensive assessment of reliability as an integral property encompassing security, fault tolerance, and compliance. They do not perform multi-criteria aggregation based on expert weights and do not provide a single reliability index.

Specialized software for reliability testing: Tools for load testing (JMeter, LoadRunner, Gatling) and stress testing, as well as for fault tolerance assessment (Chaos Monkey, Gremlin). They allow for the empirical determination of metrics such as MTBF (mean time between failures), MTTR (mean time to repair), and the point of system degradation. However, these tools provide a point-by-point, technical assessment of individual aspects under load and are not intended for comparative analysis of dissimilar ICs based on a standardized quality model.

Software quality management platforms and frameworks: Solutions that support quality standards, such as Tricentis qTest and Micro Focus ALM/Quality Center, help with test planning, defect tracking, and, to some extent, comparing results with requirements. Some support reliability-related metrics. However, they remain testing process management tools and do not include built-in algorithms for multi-criteria decision making or machine learning-based data imputation.

Academic and proprietary models for metric calculation: Mathematical models and specialized software (e.g., RELIABILITY, SHARPE, SPNP) exist for calculating probabilistic reliability metrics for complex systems based on architectural models (e.g., Markov chains, Petri nets). These tools are powerful for quantitative forecasting at the design stage, but require detailed system models, are difficult to use for existing IS, and do not take into account subjective expert assessments of business criteria [9,24].

As [Table 2](#) shows, a wide variety of tools and approaches are used in practice to assess the reliability of information systems, but each addresses only a limited range of tasks. Monitoring systems and APM are well suited for monitoring the system's state in real time, while testing tools allow for testing its behavior under load. However, neither approach provides a comprehensive understanding of reliability as an integral property. In contrast, the proposed methodology combines the requirements of the ISO/IEC 25010 standard, multi-criteria analysis, and machine learning methods, resulting in a clear and interpretable indicator of information system reliability.

Table 2: Comparative table of existing approaches and tools.

Category /Tool	Key Aspects Assessed	Advantages	Disadvantages and Limitations	Integration with Standards (ISO 25010)	Multicriteria Aggregation	Working with Incomplete Data
APM systems (Datadog, Dynatrace)	Performance, availability, real-time errors	Highly detailed, real-time monitoring, and visual dashboards.	Highly detailed, real-time monitoring, and visual dashboards.	Missing or limited to predefined dashboards.	No. Aggregation at the level of simple metrics.	No. They require a full data stream.
Testing tools (JMeter, Chaos Monkey)	Load tolerance, failure behavior.	Empirical measurement of ultimate capabilities, hypothesis testing.	The system is evaluated under artificial conditions; there is no comprehensive view of operational reliability.	Absent.	No.	No.
Quality management platforms (qTest, ALM)	Requirements coverage, number of defects, testing progress.	Link to software lifecycle, compliance tracking.	Focus on the process rather than on an integral assessment of the working system.	Partial, through custom fields and metrics.	Very limited, through user reports.	No.
Predictive reliability models (SHARPE, SPNP)	Failure probability, availability based on architectural model.	High accuracy of forecasting at the design stage, assessment of the impact of components.	They require an accurate and detailed model of the system, are not applicable to a “black box” (ready-made IS), and do not take into account expert opinions.	Absent.	No. They use probabilistic methods.	Not applicable. Works with full models.
Hybrid MCDM methods (from literature, e.g., FAHP-TOPSIS [15])	Various criteria determined by the researcher.	Taking into account qualitative assessments, compromise between criteria.	Typically implemented as scripts (MATLAB, R, Python) for a specific study, rather than as a ready-made tool. They require complete data.	Rarely. Criteria are determined by the authors.	Yes, this is their main function.	No. This is a key limitation of most works.
The proposed methodology	All aspects of reliability according to ISO 25010: availability, fault tolerance, recoverability, security, etc.	Direct link to ISO 25010. MCDM integration for aggregation. ML for missingness imputation. Interpretable reliability percentage.	Requires the involvement of experts for the AHP stage. It is a methodology requiring software implementation.	Direct and mandatory. The criteria framework is strictly defined by the standard.	Yes, enhanced. It uses an ensemble of three methods (ARAS, CoCoSo, TOPSIS) with geometric aggregation.	Yes, that’s the key difference. The built-in data recovery step is powered by XGBoost.

1.5 Comparative Analysis of Existing Methods and Software Tools for Assessing the Reliability of Information Systems

The conducted analysis of the literature allows us to clearly formulate a scientific gap: there is no comprehensive, standardized and practice-oriented methodology for assessing the reliability of information systems that would synergistically combine:

1. Normative framework (ISO/IEC 25010:2023) for structuring the task;
2. Expert-analytical procedures for weighing and aggregation;
3. Algorithms for intelligent data processing to ensure the completeness and reliability of the initial information;
4. A combination of MCDM methods that not only provides robust rankings but also generates a final, interpretable percentage reliability score.

The contribution and novelty of this study lies precisely in the development of such a comprehensive hybrid methodology. Specific scientific and practical contributions can be summarized as follows:

Conceptual and methodological contribution: For the first time, an assessment model has been proposed and formalized in which the ISO/IEC 25010:2023 standard is used as a direct and mandatory basis for constructing a hierarchy of criteria, ensuring consistency and objectivity.

Methodological contribution: A consistent integration of three disparate paradigms (standardization, MCDM, ML) into a single workflow is proposed.

Technological contribution: The Machine Learning algorithm is adapted and applied specifically to solve the imputation problem in the context of multi-criteria evaluation, which allows working with real incomplete data.

Practical contribution: The practical significance of the proposed methodology lies in its ability to comprehensively determine the reliability level of an information system, taking into account all key criteria and indicators regulated by the ISO/IEC 25010:2023 standard, including availability, fault tolerance, recoverability, performance, and security. The integration of diverse criteria into a single model provides a holistic view of information system reliability as a system property, and the final result, presented as a percentage (0%–100%), makes the resulting assessment clear, interpretable, and convenient for direct comparison of information systems with different architectures and purposes.

Contribution to verification: The combination of methods and the sensitivity analysis conducted confirm the reliability, robustness and practical applicability of the developed methodology.

This study contributes not to the creation of new algorithms, but to their innovative integration in a new applied context—a comprehensive, standardized assessment of IS reliability. The proposed methodology fills the identified gap, offering researchers and practitioners a transparent, reproducible, and sustainable tool that overcomes the fragmentation of existing approaches and delivers the final result in a format that is most convenient for management decision-making.

2 Materials and Methods

Assessing the reliability of information systems is a complex, multidimensional task, as this indicator depends on a combination of functional, operational, and quality characteristics. To ensure the objectivity and reproducibility of results, a methodological framework based on recognized international standards is necessary. This paper uses ISO/IEC 25010:2023 as such a framework, which offers a systematic approach to describing the quality of software and information systems. Its structure allows for the identification of individual criteria and subcriteria that can be formalized and used for quantitative reliability assessment.

However, using standards alone is insufficient to obtain an integrated assessment, as criteria have varying weights and may conflict with one another. To address this issue, MCDM methods are used, which allow for the consideration of weighting factors and the aggregation of disparate indicators. The study selected the ARAS, CoCoSo, and TOPSIS methods, which differ in their calculation logic but, when combined, provide a more robust and objective assessment.

Machine learning methods are also integrated into the methodology. The use of the XGBoost algorithm allows for the recovery of missing data, thereby eliminating the problem of incomplete information typical of practical research. The combination of the ISO/IEC 25010:2023 standard, MCDM methods, and machine learning forms the basis for a comprehensive and robust methodology for assessing information systems.

The ISO/IEC 25010:2023 standard provides an internationally recognized model for evaluating software and information system quality, comprising nine key criteries and over thirty subcriteries. This hierarchical model makes it possible to translate the complex concept of quality into measurable components. This approach enables the quantitative assessment of multidimensional properties of information systems and their comparison with one another.

For the purposes of this study, the most relevant criteries and subcriteries directly related to the reliability of information systems were identified from the standard's overall structure. At the first tier, the model identifies the overarching attribute of Reliability; the second tier specifies its key elements—availability, fault tolerance, and recoverability—while the third tier links each subcriteria to concrete, measurable indicators such as mean time to recovery, availability ratio, and failure frequency. Hierarchy is consistent. Weights are correctly distributed at all levels of the model. To make sure that the expert judgments are reliable, another parameter in validation is also computed—the consistency ratio (CR). If, then results can be accepted and hence used for further calculations.

To obtain an objective, integrated assessment of the reliability of information systems, it is necessary to determine the weighting factors for criteria, subcriteria, and indicators. For this purpose, the AHP is used, which allows for the consideration of expert assessments and the specifics of the subject area.

As shown in [Fig. 1](#), the model has a three-level hierarchical structure:

Level 1—criteria defined by ISO/IEC 25010:2023;

Level 2—corresponding subcriteria;

Level 3—measurable indicators that enable quantitative assessment of the subcriteria.

At each level, a pairwise comparison of elements is conducted based on their significance. Experts create judgment matrices, the values of which are taken from the Saaty scale (from 1 to 9). Local weights are calculated based on these matrices.

To ensure the correctness of the scales, the following normalization conditions are met:

the sum of the weights of all criteria is equal to 1:

$$\sum_{i=1}^n w_i = 1; \quad (1)$$

the sum of the weights of subcriteria within each criterion is equal to 1:

$$\sum_{j=1}^m w_{ij} = 1, \quad \forall i; \quad (2)$$

the sum of the weights of the indicators within each subcriterion is equal to 1:

$$\sum_{k=1}^p w_{ijk} = 1, \quad \forall i, j; \quad (3)$$

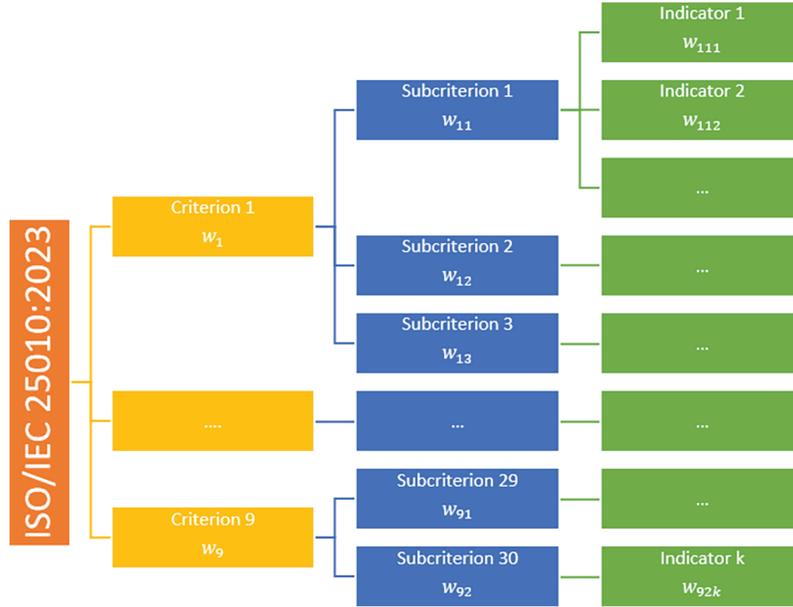


Figure 1: Three-level model for assessing the reliability of information systems.

Hierarchical consistency and the correct distribution of weights across all model levels are ensured. To ensure the reliability of expert judgments, the consistency ratio (CR) is also computed as part of the validation process. If $CR < 0.1$, the results are considered acceptable, and the resulting weights are used in subsequent calculations.

In the next step, decision matrices are formed for each criterion, with rows corresponding to alternatives and columns corresponding to indicators. The matrix elements reflect quantitative or qualitative assessments of the alternatives for each indicator.

Let there be many alternatives

$$A = \{A_1, A_2, \dots, A_m\} \quad (4)$$

and many indicators

$$C = \{C_1, C_2, \dots, C_n\} \quad (5)$$

Each alternative is assessed according to an indicator, forming a decision matrix:

$$X = (x_{ij})_{m \times n}, \quad i = 1, \dots, m; \quad j = 1, \dots, n, \quad (6)$$

where x_{ij} is the value of the i -th alternative for the j -th indicator.

The weight of each indicator is defined as the product of the weights of the criteria, subcriteria, and the indicator itself. Let w_i be the weight of the i -th criterion, w_{ij} be the weight of the j -th subcriterion related to

the given criterion, and w_{ijk} be the local weight of the k -th indicator within the corresponding subcriterion. Then the global weight of the indicator is calculated using the formula:

$$W_{ijk} = w_i \cdot w_{ij} \cdot w_{ijk}, \quad \sum_{k=1}^n W_{ijk} = 1 \tag{7}$$

Each indicator receives an integral weight that takes into account the significance of the criterion, subcriterion, and the indicator itself. The sum of the global weights of all indicators equals one, ensuring model consistency.

The next step involves checking the evaluation matrix for missing elements. If the matrix contains an element without a value:

$$\exists x_{ij} = \emptyset \rightarrow x_{ij} \text{ must be restored,} \tag{8}$$

then this element is replaced with the predicted value x'_{ij} , calculated using a machine learning model. Specifically, the XGBoost algorithm is used, which is trained on an incomplete matrix and reconstructs the missing values:

$$x'_{ij} = f_{XGBoost}(X_{\setminus j}) \tag{9}$$

$X_{\setminus j}$ is the set of observed features (other indicators) used to restore the value x_{ij} .

After data supplementation, final decision matrices are formed, which additionally include optimal values for each indicator. These values are designated as x_{0j} and are located in the first row of the matrix.

$$X = (x_{ij})_{m \times n}, \quad i = 0, \dots, m; \quad j = 1, \dots, n, \tag{10}$$

If optimal values are not present in the initial data, they are determined according to the following rule:

- if the indicator is of the max type (the higher the better), then the highest value among all alternatives is considered optimal;
- if the indicator is of the min type (the smaller the better), then the smallest value is optimal.

Formally:

$$x_{0j} = \begin{cases} \max_{1..i} x_{ij}, & \text{if the indicator is of the "max" type;} \\ \min_{1..i} x_{ij}^*, & \text{if the indicator is of the "min" type.} \end{cases} \tag{11}$$

2.1 Application of the ARAS (Additive Ratio Assessment) Method

The ARAS method is used to obtain an integrated assessment of alternatives based on their relative proximity to the optimal solution. The basic idea of the method is that each alternative is normalized relative to the sum of all values for each indicator, after which a weighted sum of the normalized values is determined. The higher the value of the integrated indicator, the more preferable the alternative.

Each element of the matrix is normalized relative to the sum of the values for the corresponding indicator:

$$r_{ij} = \frac{x_{ij}}{\sum_{i=0}^m x_{ij}} \text{ (for the max indicator)}$$

$$r_{ij} = \frac{1/x_{ij}}{\sum_{i=0}^m 1/x_{ij}} \text{ (for the min indicator)} \quad (12)$$

Taking into account the weights of the indicators w_j , a weighted normalized matrix is calculated:

$$v_{ij} = w_j \cdot r_{ij} \quad (13)$$

For each alternative, the total priority is determined:

$$S_i = \sum_{j=1}^n v_{ij} \quad (14)$$

The utility of each alternative is calculated as the ratio of its total value to the value of the optimal alternative:

$$K_i = \frac{S_i}{S_0} \quad (15)$$

where S_0 is the value for the ideal alternative.

The alternatives are ordered in descending order of K_i . The best alternative has the highest utility value.

- Ease of computation and clarity in interpreting the results;
- Obtaining a visual assessment of the utility of each alternative compared to the optimal one;
- Widely applicable in economics, management, engineering, and information systems evaluation.

The ARAS method is an effective tool for multi-criteria analysis, allowing for informed decision-making based on a comprehensive assessment of alternatives.

2.2 Application of the CoCoSo (Combined Compromise Solution) Method

The CoCoSo method is a modern approach to multicriteria analysis that combines the principles of weighted sum and weighted tradeoffs to obtain an integrated assessment of alternatives. Each alternative in the CoCoSo method is normalized by each indicator, after which aggregate indices are calculated that take into account both average values and relative priorities. The final ranking is based on a combined utility measure, which allows for the balance between benefit and cost indicators to be considered.

Normalization of indicators (taking into account the maximization or minimization of criteria):

$$r_{ij} = \frac{x_{ij}}{\max_i x_{ij}} \text{ (for max criteria)}$$

$$r_{ij} = \frac{\min_i x_{ij}}{x_{ij}} \text{ (for min criteria)} \quad (16)$$

Weighting of normalized values

$$v_{ij} = w_j \cdot r_{ij} \quad (17)$$

where w_j is the weight of the j -th criterion.

Calculation of additive and multiplicative aggregate functions:

Summary function:

$$S_i = \sum_{j=1}^n v_{ij} \tag{18}$$

The function produced:

$$P_i = \prod_{j=1}^n v_{ij}^{w_j} \tag{19}$$

Based on these, combined indices are formed:

$$Q_i^{(1)} = \frac{S_i + P_i}{2} \tag{20}$$

$$Q_i^{(2)} = \frac{S_i}{\max(S)} + \frac{P_i}{\max(P)} \tag{21}$$

$$Q_i^{(3)} = \frac{\lambda S_i + (1 - \lambda) P_i}{\max_i (\lambda S_i + (1 - \lambda) P_i)} \tag{22}$$

$$K_i = \frac{1}{3} (Q_i^{(1)} + Q_i^{(2)} + Q_i^{(3)}) \tag{23}$$

$\lambda \in [0, 1]$ —equilibrium coefficient.

The alternatives are ranked by their K_i values in descending order. The alternative with the highest value is considered the best.

The advantages of using the CoCoSo method show high stability to weight changes, a compromise between an additive and multiplicative schemes, and universal application in engineering, economics, and reliability assessment of information systems.

2.3 Application of the TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) Method

Regarded as one of the leading tools in multicriteria analysis, the TOPSIS method operates on the notion of choosing the alternative that exhibits the smallest distance to the ideal solution while remaining the farthest from the anti-ideal reference point. Thus, it first constructs a normalized, weighted decision matrix and then defines ideal and anti-ideal vectors corresponding to each indicator. For each alternative, the distances to these reference points are calculated, after which a proximity coefficient is calculated. Alternatives are ranked by the magnitude of this coefficient: the higher the coefficient, the more preferable the alternative.

Vector normalization is applied:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}} \tag{24}$$

after which the normalized values are weighed:

$$v_{ij} = w_j \cdot r_{ij} \tag{25}$$

Then the ideal solutions are determined:

- Positive Ideal Solution (PIS):

$$V^+ = \{v_1^+, v_2^+, \dots, v_n^+\} \quad (26)$$

where

$$\begin{aligned} v_i^+ &= \max_j v_{ij}, \text{ if the indicator is of the "max" type} \\ v_i^+ &= \min_j v_{ij}, \text{ if the indicator is of the "min" type} \end{aligned} \quad (27)$$

- Negative-ideal solution (NIS):

$$V^- = \{v_1^-, v_2^-, \dots, v_n^-\} \quad (28)$$

where

$$\begin{aligned} v_i^- &= \min_j v_{ij}, \text{ if the indicator is of the "max" type} \\ v_i^- &= \max_j v_{ij}, \text{ if the indicator is of the "min" type} \end{aligned} \quad (29)$$

The distance of the alternative V_i to the positive and negative solutions is calculated using the Euclidean metric:

$$\begin{aligned} S_i^+ &= \sqrt{\sum_{j=1}^n (v_{ij} - v_i^+)^2} \\ S_i^- &= \sqrt{\sum_{j=1}^n (v_{ij} - v_i^-)^2} \end{aligned} \quad (30)$$

The relative closeness of an alternative to the ideal solution is defined as

$$K_i = \frac{S_i^-}{S_i^+ + S_i^-}, 0 \leq K_i \leq 1 \quad (31)$$

Alternatives are sorted in descending order of K_i values. The larger the K_i value, the closer the alternative is to the ideal solution.

Advantages of the TOPSIS method:

- Simplicity and clarity—the method is distinguished by clear mathematical procedures and easily interpretable results;
- Taking into account ideal and anti-ideal solutions—allows you to objectively evaluate alternatives, focusing on both the best and the worst solution;
- The incorporation of criterion weights provides flexibility in decision-making by accounting for the varying importance of individual indicators.

At the concluding stage, the measured values of the indicators are evaluated in relation to their optimal counterparts within the decision matrix. Each alternative's divergence from the optimal value is determined, which makes it possible to represent its reliability degree as a percentage from 0% to 100%.

Next, an integrated reliability assessment is formed by aggregating the results obtained using various multicriteria analysis methods (ARAS, CoCoSo, TOPSIS). To ensure a balanced final metric, the geometric mean of the integrated indicators is calculated:

$$R_i = \left(K^{(ARAS)} \cdot K^{(CoCoSo)} \cdot K^{(TOPSIS)} \right)^{\frac{1}{3}} \tag{32}$$

where $K^{(ARAS)}$, $K^{(CoCoSo)}$, $K^{(TOPSIS)}$ are normalized integral estimates of the i -th alternative obtained by the corresponding methods.

Based on the R_i values, a composite ranking of alternatives is formed. The higher the R_i value, the closer the system is to optimal performance indicators and the greater its reliability.

As illustrated in Fig. 2, the proposed reliability assessment framework outlines the complete workflow of the hybrid methodology that integrates the ISO/IEC 25010:2023 quality model, machine learning, and MCDM techniques. The process begins with the selection of indicators and data entry, followed by a validation step to check the completeness of data. In order to have a fully completed dataset, missing entries are filled in by outcome estimations through the XGBoost algorithm. Expert judgment using AHP is applied in determining weights for criteria, subcriteria, and indicators. The calculated weights are then applied within ARAS, CoCoSo, and TOPSIS methods that return integrated reliability scores. The final reliability index is taken as a geometric mean value of these results—this single number summary neatly encapsulates information system reliability.

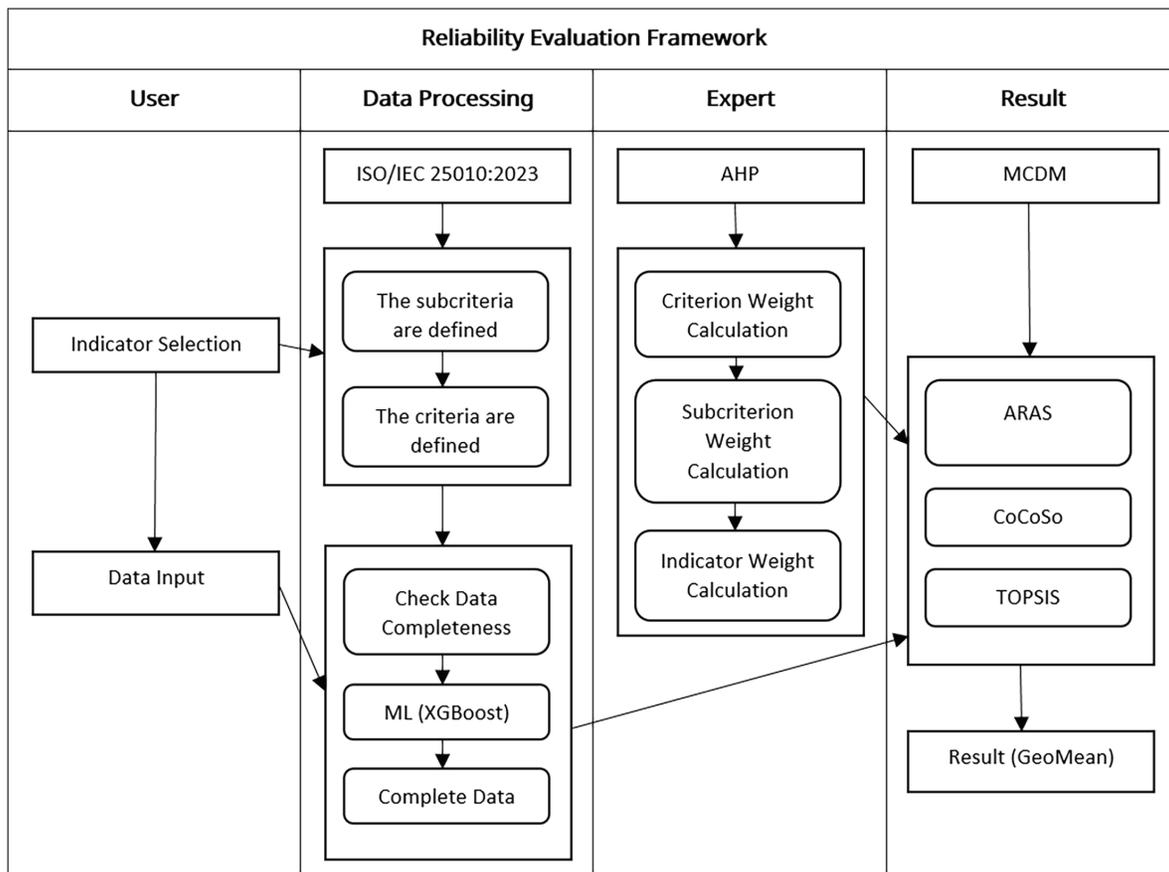


Figure 2: Reliability evaluation framework for information systems.

In summary, this methodology enables and combines the use of the most recent international standard ISO/IEC 25010:2023, multicriteria analysis methods (ARAS, CoCoSo, TOPSIS), AHP for weighting determination, and machine learning algorithms (XGBoost) for imputation of missing values. Through the integration of measurable indicators and weighted significance, the hybrid model delivers a thorough and impartial assessment of information system reliability, expressed as an integrated percentage value. The final composite score, obtained via the geometric mean across various MCDM approaches, guarantees stable and comprehensible results, confirming the methodology's suitability for applied and scholarly use alike.

3 Results

For our research, we needed real data. To this end, we contacted several IT companies, who agreed to assist us and provided information about the information systems they supported. The companies strictly insisted on confidentiality: we cannot disclose the names of the systems themselves or the companies. We were provided only with system descriptions and measurement results for the indicators necessary for assessing reliability.

Based on this data, we created a dataset that we use for machine learning and multicriteria analysis. As an example experiment, we included data from five different information systems. Their key characteristics are presented in the table below (Table 3).

Table 3: Characteristics of the studied information systems.

Information System	Characteristics
Information System No. 1	Hardware: 2 × Intel Xeon Silver 4210 CPU, 128 GB RAM, 4 TB SSD, 10 Gbps LAN. Operating system: Ubuntu Server 22.04 LTS. Database: PostgreSQL 14. System architecture: Microservices. Programming languages and frameworks: Java (Spring Boot), JavaScript (React). Interface types: Web interface, mobile application. System purpose: student registration, class schedule management, exam results storage. Integration options: LMS (Moodle), HR system, library collection via API.
Information System No. 2	Hardware: Intel Xeon Gold 6226R, 64 GB RAM, 2 TB SSD, 5 Gbps LAN. Operating system: Windows Server 2019. Database: MS SQL Server 2019. System architecture: Monolithic. Programming languages and frameworks: C# (.NET Core), Angular. Interface types: Web interface, desktop application (Windows). System purpose: maintaining patient medical records, scheduling appointments, storing lab results. Integration capabilities: data exchange with medical equipment (HL7/FHIR), e-Gov API.
Information System No. 3	Hardware: 4 × Intel Xeon Platinum 8358, 256 GB RAM, 8 TB SSD RAID, hardware security module (HSM). Operating system: Red Hat Enterprise Linux 9. Database: Oracle Database 19c. System architecture: Service-Oriented Architecture (SOA). Programming languages and frameworks: Java (Spring), Kotlin, Vue.js. Interface types: Web interface, mobile application. System purpose: payments, online banking. Integration options: national payment system, SWIFT.

(Continued)

Table 3 (continued)

Information System	Characteristics
Information System No. 4	Hardware: Intel Xeon E5-2680, 32 GB RAM, 1 TB SSD, 1 Gbps LAN. Operating system: Debian 11. Database: MySQL 8. System architecture: Microservices. Programming languages and frameworks: Python (Django, FastAPI), JavaScript (Vue.js). Interface types: Web interface, mobile application (Android). System purpose: crop monitoring, agricultural warehouse management, climate data analysis. Integration capabilities: drone data reception via API, satellite data connection.
Information System No. 5	Hardware: 2 × Intel Xeon Gold 5318, 128 GB RAM, 6 TB SSD, Dual 10 Gbps LAN. Operating system: Windows Server 2022 Datacenter. Database: MongoDB (for documents), PostgreSQL (for relational data). System architecture: Hybrid (microservices + REST API). Programming languages and frameworks: Java (Spring Boot), Node.js, React. Interface types: Web interface, mobile application. System purpose: Providing services to the population (registration, utility bills, filing applications). Integration options: e-Gov, banking payment systems, SMS/e-mail notification gateways.

To provide a clearer presentation of the experimental data, the above-described information systems will be considered in a simplified variant. To analyze their reliability, a set of key indicators is used, which most fully and systematically characterize the system features. The list of indicators is provided in [Table 4](#). This makes it possible to present the data compactly, concurrently performing multi-criteria analysis and visualization of the results.

Table 4: List of indicators for analyzing the reliability of information systems.

No.	Indicator Code	Indicator	How is it Measured?
1	X1	Availability	The ratio of system operation time to total observation time
2	X2	Mean time to recovery (MTTR)	Average time to recover from failure
3	X3	Mean time between failures (MTBF)	Ratio of total uptime to failure rate
4	X4	Failure rate	Number of failures per unit of operating time
5	X5	System response time	Average system response time to a user request
6	X6	CPU usage	Average CPU load under typical workload
7	X7	RAM usage	Average RAM usage while running operations
8	X8	Percentage of functional requirements fulfilled	Number of implemented functions/number of required functions
9	X9	Accuracy of data processing	The proportion of correct data processing results
10	X10	Access security level (percentage of successful attacks)	Number of successful attacks/total number of attack attempts

The AHP pairwise comparison method was used to calculate the weights. The evaluation took into account the average values obtained from 25 experts. The panel of 25 experts included IT system architects (8), software engineers (7), information security specialists (5), and academic researchers in information systems (5). All experts had at least 7 years of professional experience and prior participation in system evaluation or quality assurance projects. The consistency ratio (CR) of all matrices was below 0.1, indicating acceptable agreement among the experts.

Based on this data, the criteria, subcriteria, and indicators were prioritized, after which the values were normalized. The final results are summarized in Table 5, and a graphical representation of the structure and distribution of the weights is shown in Fig. 3.

Table 5: Results of determining weights using the AHP method.

Criterion	Weight of the Criterion	Subcriterion	Subcriterion Weight	Indicator	Weight of the Indicator	Normalized Value of the Indicator
Reliability	0.3269	Availability	0.2967	X1	1	0.097
		Recoverability	0.23	X2	1	0.0752
		Maturity	0.2233	X3	1	0.0730
		Fault tolerance	0.25	X4	1	0.0817
Performance efficiency	0.2808	Time behaviour	0.5143	X5	1	0.1444
		Resource utilization	0.4857	X6	0.5143	0.0701
Functional suitability	0.2269	Functional completeness	0.4747	X8	1	0.1077
		Functional correctness	0.5253	X9	1	0.1192
Security	0.1654	Integrity	1	X10	1	0.1654
General	1					1

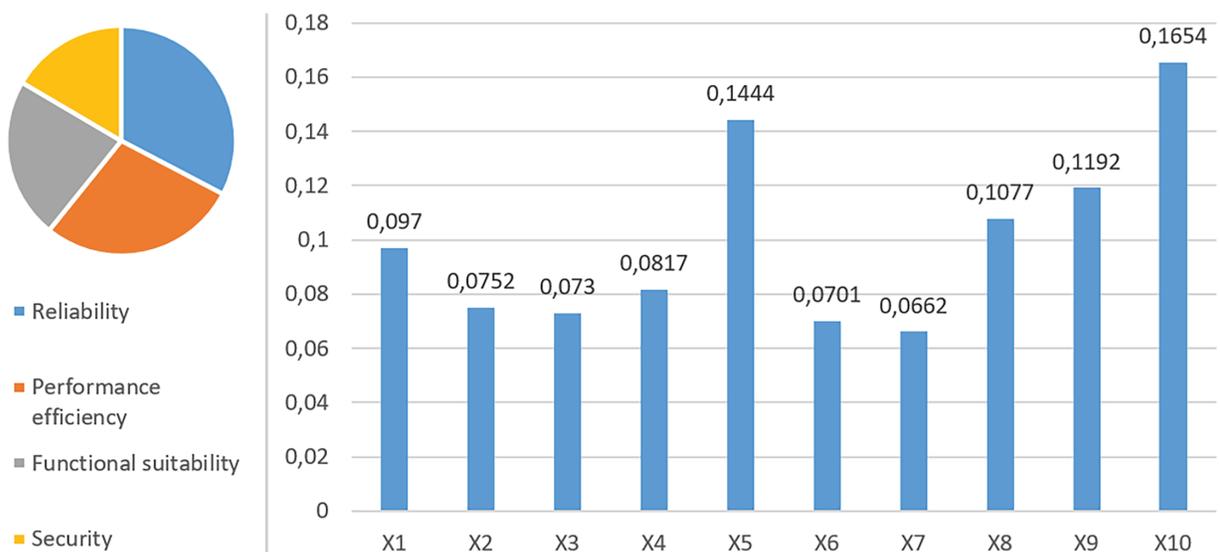


Figure 3: Graphical representation of the distribution of weights of criteria and indicators.

Table 6 presents the indicator values for the five experimental information systems, along with their optimal values and optimal directions. A sixth system was also created, similar to the first. Several gaps were intentionally left in its data, which will later be reconstructed using machine learning methods to demonstrate the process of processing incomplete data and compare the obtained results.

Table 6: Values of indicators for experimental information systems.

IS	Criteria									
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
Units of measurement	Share (0–1)	Hours	Hours	Refusal/hour	Seconds	%	%	%	Share (0–1)	Share (0–1)
Direction of optimization	Max	Min	Max	Min	Min	Min	Min	Max	Max	Min
Weight of the criterion	0.097	0.0752	0.0730	0.0817	0.1444	0.0701	0.0662	0.1077	0.1192	0.1654
Optimal value	≥0.999	≤1	≥1000	≤0.01	≤2	≤70	≤75	100	≥0.99	≤0.01
	0.999	0.7	1200	0.008	1.5	65	69	100	0.993	0.006
IS No. 1	0.995	0.8	1200	0.008	1.5	65	70	99	0.992	0.008
IS No. 2	0.992	0.9	950	0.010	1.7	68	74	97	0.988	0.007
IS No. 3	0.989	1.1	1100	0.012	1.8	72	69	98	0.991	0.008 (ML)
IS No. 4	0.990	1.00 (ML)	980	0.009	2.1	75	76 (ML)	96	0.985	0.009
IS No. 5	0.993 (ML)	0.7	1050	0.011	1.6	66	73	100	0.993	0.006
IS No. 6	0.995	0.82 (ML)	1200	0.008	1.5	65	70 (ML)	99	0.992 (ML)	0.008

For indicators with an optimal “max” direction, if the actual value is higher than the specified optimal value, then the highest of the presented values is taken as a reference value. In case it is lower, the original optimal value remains. For indicators with a “min” optimization direction, the same logic is followed: if the measured value is smaller than the lower limit, that minimum is used as the reference, while higher values preserve the original optimal setting. This approach allows for the specific characteristics of indicators to be accurately taken into account during comparative analysis and the formation of a summary assessment.

As shown in Fig. 4, a final matrix was generated, reflecting the impact of each indicator on the overall reliability of the systems. This data presentation format allows for comparison not only of alternatives but also of the optimal value, highlighting the strengths and weaknesses of each information system. The obtained results served as the basis for further calculation of integrated indicators using the ARAS method and their visualization in the form of a diagram.

The final results calculated using the CoCoSo method are presented in Fig. 5. These values represent composite scores of all indicators and permit relative analysis among the six information systems.

Fig. 6 shows the results of TOPSIS. While comparing alternatives, the optimization direction for each criterion is taken into consideration: for indicators with a “max” direction, higher values are considered to indicate better system quality; for indicators with a ‘min’ direction, the minimum value is considered as good. This makes the comparison of alternatives fair and exposes, in a more objective manner, the strong and weak points of the information systems under study.

ARAS

IS	Criteria									
	X1 Ⓢ	X2 Ⓢ	X3 Ⓢ	X4 Ⓢ	X5 Ⓢ	X6 Ⓢ	X7 Ⓢ	X8 Ⓢ	X9 Ⓢ	X10 Ⓢ
Optimal value	0.0139	0.0129	0.0114	0.0134	0.0227	0.0105	0.0098	0.0156	0.0171	0.0286
IS №1	0.0139	0.0113	0.0114	0.0134	0.0227	0.0105	0.0097	0.0155	0.0171	0.0215
IS №2	0.0138	0.0100	0.0090	0.0107	0.0200	0.0100	0.0091	0.0152	0.0170	0.0246
IS №3	0.0138	0.0082	0.0105	0.0090	0.0189	0.0094	0.0098	0.0153	0.0170	0.0215
IS №4	0.0138	0.0090	0.0093	0.0119	0.0162	0.0091	0.0089	0.0150	0.0169	0.0191
IS №5	0.0139	0.0129	0.0100	0.0098	0.0213	0.0103	0.0093	0.0156	0.0171	0.0286
IS №6	0.0139	0.0110	0.0114	0.0134	0.0227	0.0105	0.0097	0.0155	0.0171	0.0215

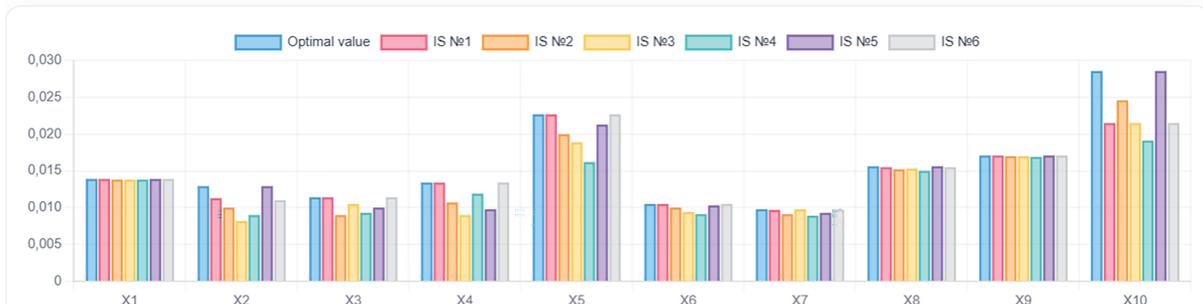


Figure 4: Results of applying the ARAS method.

CoCoSo

IS	Criteria									
	X1 Ⓢ	X2 Ⓢ	X3 Ⓢ	X4 Ⓢ	X5 Ⓢ	X6 Ⓢ	X7 Ⓢ	X8 Ⓢ	X9 Ⓢ	X10 Ⓢ
Optimal value	0.0970	0.0752	0.0730	0.0817	0.1444	0.0702	0.0662	0.1077	0.1192	0.1654
IS №1	0.0966	0.0658	0.0730	0.0817	0.1444	0.0702	0.0653	0.1066	0.1191	0.1241
IS №2	0.0963	0.0585	0.0578	0.0654	0.1274	0.0671	0.0617	0.1045	0.1186	0.1418
IS №3	0.0960	0.0479	0.0669	0.0545	0.1203	0.0634	0.0662	0.1055	0.1190	0.1241
IS №4	0.0961	0.0526	0.0596	0.0726	0.1031	0.0608	0.0601	0.1034	0.1182	0.1103
IS №5	0.0964	0.0752	0.0639	0.0594	0.1354	0.0691	0.0626	0.1077	0.1192	0.1654
IS №6	0.0966	0.0642	0.0730	0.0817	0.1444	0.0702	0.0653	0.1066	0.1191	0.1241

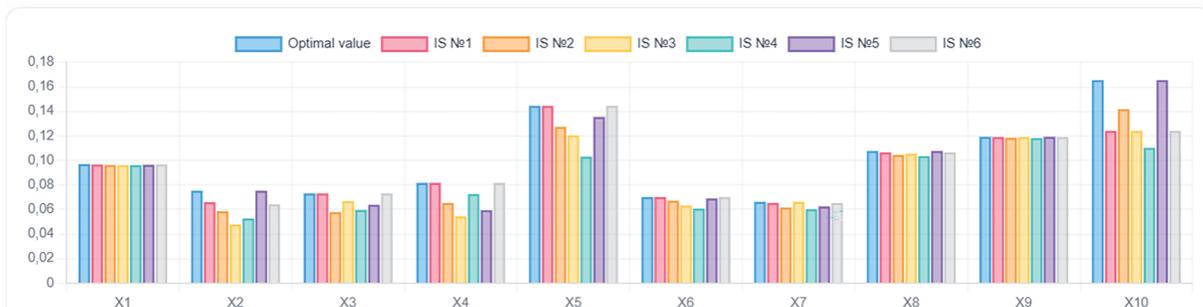


Figure 5: Results of applying the CoCoSo method.

TOPSIS

IS	Criteria									
	X1 Ⓢ	X2 Ⓢ	X3 Ⓢ	X4 Ⓢ	X5 Ⓢ	X6 Ⓢ	X7 Ⓢ	X8 Ⓢ	X9 Ⓢ	X10 Ⓢ
	max	min	max	min	min	min	min	max	max	min
Optimal value	0.0369	0.0228	0.0301	0.0259	0.0486	0.0253	0.0241	0.0414	0.0452	0.0500
IS №1	0.0367	0.0261	0.0301	0.0259	0.0486	0.0253	0.0245	0.0409	0.0451	0.0667
IS №2	0.0366	0.0294	0.0238	0.0323	0.0551	0.0265	0.0259	0.0401	0.0449	0.0583
IS №3	0.0365	0.0359	0.0276	0.0388	0.0583	0.0281	0.0241	0.0405	0.0451	0.0667
IS №4	0.0365	0.0326	0.0245	0.0291	0.0681	0.0292	0.0266	0.0397	0.0448	0.0750
IS №5	0.0367	0.0228	0.0263	0.0356	0.0519	0.0257	0.0255	0.0414	0.0452	0.0500
IS №6	0.0367	0.0268	0.0301	0.0259	0.0486	0.0253	0.0245	0.0409	0.0451	0.0667

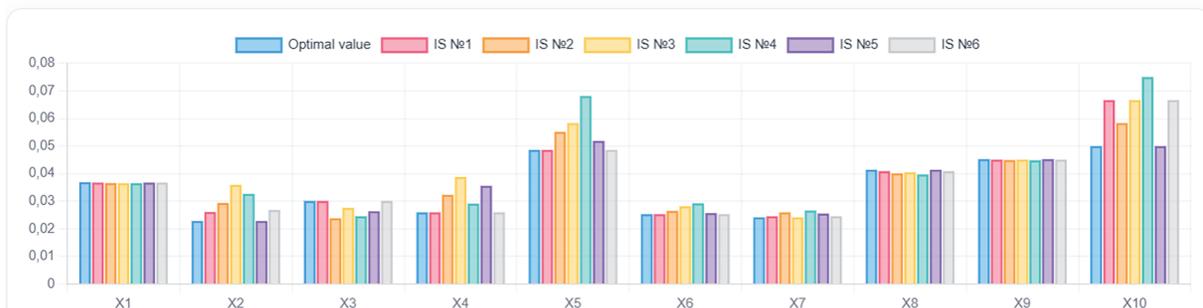


Figure 6: Results of applying the TOPSIS method.

4 Discussion

Reliability assessment was conducted using three multicriteria analysis methods: ARAS, CoCoSo, and TOPSIS. The results showed that, according to the ARAS and CoCoSo methods, all six systems under study had values close to optimal, indicating a relatively high level of reliability. IS No. 1 and No. 5 demonstrated the most consistent results, while IS No. 3 and No. 4 showed relatively low values (Fig. 7).

To interpret the obtained reliability percentage values, a scale based on recommendations from international software quality standards was used. This scale allows for the classification of information systems by reliability levels and the development of practical recommendations (Table 7).

An extra sensitivity test was carried out to check and confirm the stability of our hybrid methodology under different operational conditions. The weights of the indicators, which were set through the AHP method, have been slightly modified by $\pm 10\%$. The total composite scores were then recalculated again in TOPSIS and CoCoSo methods. It appeared from the analysis that small changes in weight did not disturb the final ranking of information systems. Very slight changes have been observed only at a few borderline cases mostly in “Security” and “Response Time” metrics. Therefore, it can be considered as an evident result proving robustness and stability of model when weighting configuration changes within small limits.

General

IS	Method			
	ARAS	CoCoSo	TOPSIS	Average ^①
Optimal value	1.0000	1.0000	1.0000	1.0000
IS №1	0.9414	0.9452	0.6205	0.8204
IS №2	0.8944	0.8981	0.5979	0.7832
IS №3	0.8554	0.8607	0.3374	0.6286
IS №4	0.8290	0.8340	0.2315	0.5429
IS №5	0.9533	0.9534	0.7500	0.8801
IS №6	0.9396	0.9435	0.6161	0.8174

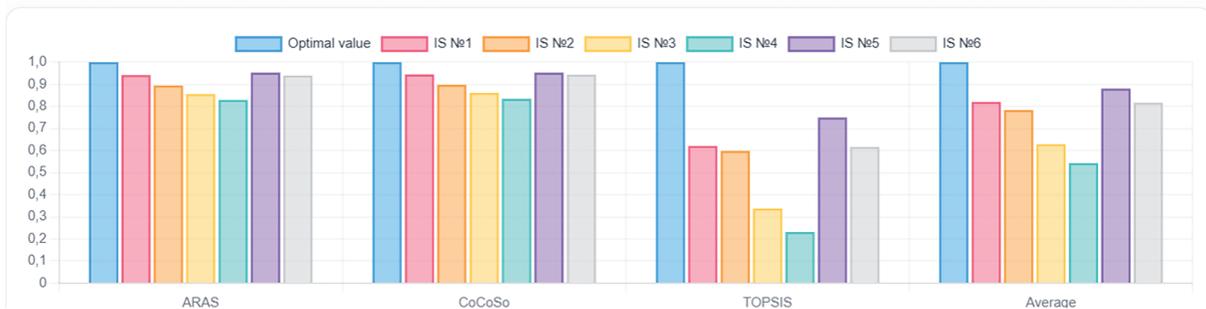


Figure 7: Final results of the information systems reliability assessment.

Table 7: Scale for interpreting percentage values of reliability of information systems.

Scale	Description
0%–20%—extremely low level of reliability	As the system falls short of minimum operational standards and demonstrates a high likelihood of failure, significant architectural revisions, resource optimization, and the adoption of redundancy strategies are necessary.
20%–40%—low level of reliability	The system is working but it is prone to failures and high load conditions. Improvements in fault tolerance and security and monitoring and diagnostic mechanisms are needed.
40%–60%—average reliability level	The system meets basic requirements but still shows significant limitations. Gradual performance improvement and fine-tuning of functionality are recommended to minimize downtime and the number of errors.
60%–80%—high level of reliability	The system meets the requirements of users and accepted benchmarks. Occasionally, a weakness may be observed during high-load periods. To maintain its current state, constant tuning and optimization should be conducted.
80%–100%—very high level of reliability	The system is performing at its best, or almost best, strictly within the definitions of ISO/IEC 25010:2023, thereby making it also a <i>de facto</i> benchmark standard. To keep up with this supervision by bringing something new and improved constantly.

5 Conclusions

This paper proposes and explores a hybrid methodology for assessing the reliability of information systems based on the combined use of ISO/IEC 25010:2023, multicriteria analysis methods (ARAS, CoCoSo, and TOPSIS), the Analytical Hierarchy Process, and machine learning algorithms. The primary goal of the study was to overcome the fragmentation of existing approaches and develop a holistic, quantitatively interpretable reliability assessment applicable to real-world information system operations.

ISO/IEC 25010:2023 was chosen as the conceptual basis for the methodology, allowing reliability to be viewed not as a standalone indicator, but as a systemic property encompassing availability, fault tolerance, maturity, and recoverability. The standard's hierarchical structure ensured a logical decomposition of requirements and linked abstract quality characteristics to specific measurable indicators. The use of the AHP method made it possible to incorporate expert knowledge and accurately determine the relative importance of criteria and indicators. A significant practical aspect of the work was addressing the problem of data incompleteness, which is typical when assessing real-world information systems. To this end, machine learning was integrated into the methodology for missing value imputation. Experiments demonstrated that using machine learning at the data preparation stage improves the completeness and consistency of the initial data without distorting the final assessment results.

To generate the final reliability assessment, three different multicriteria analysis methods—ARAS, CoCoSo, and TOPSIS, were used. Their combined use reduced the impact of the limitations of individual methods and increased the robustness of the results. The resulting assessments demonstrated high consistency between the methods, and sensitivity analysis confirmed the stability of the final conclusions when changing criterion weights within acceptable limits.

The practical value of the proposed methodology lies in its ability to comprehensively determine the reliability level of an information system, taking into account all the key criteria regulated by the ISO/IEC 25010:2023 standard. The final assessment, presented in a visual and interpretable form, facilitates the analysis of system status, the identification of weaknesses, and the comparison of information systems with different architectures and purposes. This makes the developed approach a convenient tool for supporting decisions in quality management, modernization, and the development of digital systems.

Overall, the proposed methodology represents a reproducible and practice-oriented tool for assessing the reliability of information systems. This work does not aim to create new algorithms, but rather demonstrates the sound integration of existing methods into a single model adapted to modern requirements and real-world operating conditions, which determines its scientific and practical significance.

Acknowledgement: Not applicable.

Funding Statement: The work was carried out with the financial support of the Ministry of Science and Higher Education of the Republic of Kazakhstan, grant No. AP25793823.

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Nurbek Sissenov and Dina Satybalдина; methodology, Nurbek Sissenov and Dina Satybalдина; software, Nurbek Sissenov and Gulden Ulyukova; validation, Nurbek Sissenov and Gulden Ulyukova; formal analysis, Nurbek Sissenov; investigation, Nurbek Sissenov and Gulden Ulyukova; resources, Gulden Ulyukova; data curation, Nurbek Sissenov, Dina Satybalдина and Nikolaj Goranin; writing—original draft preparation, Nurbek Sissenov and Gulden Ulyukova; writing—review and editing, Dina Satybalдина and Nikolaj Goranin; visualization, Nurbek Sissenov; supervision, Nurbek Sissenov and Dina Satybalдина; project administration, Nurbek Sissenov and Dina Satybalдина; funding acquisition, Nurbek Sissenov and Dina Satybalдина. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: Data available within the article.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

AHP	Analytic Hierarchy Process
ARAS	Additive Ratio Assessment
CoCoSo	COmbined COmpromise Solution
TOPSIS	Technique for Order of Preference by Similarity to Ideal Solution
XGBoost	Extreme Gradient Boosting
MCDM	MultiCriteria Modeling Methods
IS	Information System
ML	Machine Learning
CR	Consistency Ratio

References

1. Carpitella S, Carpitella F, Certa A, Brentan B, Izquierdo J. Advanced risk management software for multi-criteria decision-making in uncertain scenarios. *Comput Appl Math*. 2025;44(7):379. doi:10.1007/s40314-025-03338-0.
2. Tworek K. Reliability of information systems in organization in the context of banking sector: empirical study from Poland. *Cogent Bus Manage*. 2018;5(1):1522752. doi:10.1080/23311975.2018.1522752.
3. Dubey SK, Jasra B. Reliability assessment of component based software systems using fuzzy and ANFIS techniques. *Int J Syst Assur Eng Manage*. 2017;8(2):1319–26. doi:10.1007/s13198-017-0602-z.
4. Foidl H, Felderer M. Integrating software quality models into risk-based testing. *Softw Qual J*. 2018;26(2):809–47. doi:10.1007/s11219-016-9345-3.
5. Grote T, Genin K, Sullivan E. Reliability in machine learning. *Philos Compass*. 2024;19(5):e12974. doi:10.1111/phc3.12974.
6. Vinodh S, Swarnakar V. Lean Six Sigma project selection using hybrid approach based on fuzzy DEMATEL–ANP–TOPSIS. *Int J Lean Six Sigma*. 2015;6(4):313–38. doi:10.1108/ijlss-12-2014-0041.
7. Budiman E, Wati M, Widians JA, Puspitasari N, Firdaus M, Alameka F. ISO/IEC 9126 quality model for evaluation of student academic portal. In: proceedings of 2nd International Conference on Business, Finance, Management and Economics (BIZFAME); 2023 Jun 22–23; Kedah, Malaysia. doi:10.21834/e-bpj.v8iSII5.5104.
8. ISO/IEC 25010:2023. Systems and software engineering—systems and software Quality Requirements and Evaluation (SQuARE)—product quality model. Geneva, Switzerland: International Organization for Standardization; 2023.
9. Tamura Y, Kawakami M, Yamada S. Reliability modeling and analysis for open source cloud computing. *Proc Inst Mech Eng Part O J Risk Reliab*. 2013;227(2):179–86. doi:10.1177/1748006x12475110.
10. Gobov D, Zuieva O. Software quality attributes in requirements engineering. *Int J Inf Technol Comput Sci*. 2025;17(4):38–48. doi:10.5815/ijitcs.2025.04.04.
11. Raut RD, Bhasin HV, Kamble SS. Evaluation of supplier selection criteria by combination of AHP and fuzzy DEMATEL method. *Int J Bus Innov Res*. 2011;5(4):359. doi:10.1504/ijbir.2011.041056.
12. Băcescu Ene GV, Stoia MA, Cojocaru C, Todea DA. SMART multi-criteria decision analysis (MCDA)-one of the keys to future pandemic strategies. *J Clin Med*. 2025;14(6):1943. doi:10.3390/jcm14061943.
13. Yazdani M, Zarate P, Kazimieras Zavadskas E, Turskis Z. A combined compromise solution (CoCoSo) method for multi-criteria decision-making problems. *Manag Decis*. 2019;57(9):2501–19. doi:10.1108/md-05-2017-0458.
14. Ershadi MJMJ, Forouzandeh M. Information security risk management of research information systems: a hybrid approach of fuzzy FMEA, AHP, TOPSIS and Shannon entropy. *J Digit Inf Manage*. 2019;17(6):321. doi:10.6025/jdim/2019/17/6/321-336.

15. Hemmati N, Rahiminezhad Galankashi M, Imani DM, Mokhtab Rafiei F. An integrated fuzzy-AHP and TOPSIS approach for maintenance policy selection. *Int J Qual Reliab Manage.* 2020;9(10):1275–99. doi:10.1108/ijqrm-10-2018-0283.
16. Pamucar D, Ecer F, Gligorić Z, Gligorić M, Deveci M. A novel WENSLO and ALWAS multicriteria methodology and its application to green growth performance evaluation. *IEEE Trans Eng Manage.* 2024;71:9510–25. doi:10.1109/TEM.2023.3321697.
17. Mardani A, Jusoh A, Zavadskas EK. Fuzzy multiple criteria decision-making techniques and applications—Two decades review from 1994 to 2014. *Expert Syst Appl.* 2015;42(8):4126–48. doi:10.1016/j.eswa.2015.01.003.
18. Kazimieras Zavadskas E, Kazimieras Zavadskas E, Antucheviciene J, Adeli H, Turskis Z, Adeli H. Hybrid multiple criteria decision making methods: a review of applications in engineering. *Sci Iran.* 2016;23(1):1–20. doi:10.24200/sci.2016.2093.
19. Demir G, Chatterjee P, Pamucar D. Sensitivity analysis in multi-criteria decision making: a state-of-the-art research perspective using bibliometric analysis. *Expert Syst Appl.* 2024;237:121660. doi:10.1016/j.eswa.2023.121660.
20. Saraygord Afshari S, Enayatollahi F, Xu X, Liang X. Machine learning-based methods in structural reliability analysis: a review. *Reliab Eng Syst Saf.* 2022;219:108223. doi:10.1016/j.res.2021.108223.
21. Chatterjee K, Zavadskas EK, Tamošaitienė J, Adhikary K, Kar S. A hybrid MCDM technique for risk management in construction projects. *Symmetry.* 2018;10(2):46. doi:10.3390/sym10020046.
22. Mostafa AM. An MCDM approach for cloud computing service selection based on best-only method. *IEEE Access.* 2021;9:155072–86. doi:10.1109/ACCESS.2021.3129716.
23. Grimmelhuijsen S. Explaining why the computer says no: algorithmic transparency affects the perceived trustworthiness of automated decision-making. *Public Adm Rev.* 2023;83(2):241–62. doi:10.1111/puar.13483.
24. Boranbayev A, Boranbayev S, Sissenov N, Goranin N. Method and software system for assessing the reliability of information systems. *J Theor Appl Inf Technol.* 2021;99(19):4436–47. doi:10.1007/978-3-030-89912-7_38.
25. Jaiswal A, Malhotra R. Software reliability prediction using machine learning techniques. *Int J Syst Assur Eng Manage.* 2018;9(1):230–44. doi:10.1007/s13198-016-0543-y.
26. Tamascelli N, Campari A, Parhizkar T, Paltrinieri N. Artificial Intelligence for safety and reliability: a descriptive, bibliometric and interpretative review on machine learning. *J Loss Prev Process Ind.* 2024;90:105343. doi:10.1016/j.jlp.2024.105343.
27. Thomas T, Rajabi E. A systematic review of machine learning-based missing value imputation techniques. *Data Technol Appl.* 2021;55(4):558–85. doi:10.1108/dta-12-2020-0298.
28. Bakurova A, Bilyi V. Development of an approach for predicting the cost of damaged infrastructure recovery with microservice implementation. *Technol Audit Prod Reserv.* 2025;5(2(85)):33–9. doi:10.15587/2706-5448.2025.339773.
29. Levitin G, Finkelstein M, Xiang Y. Optimal inspections and mission abort policies for multistate systems. *Reliab Eng Syst Saf.* 2021;214:107700. doi:10.1016/j.res.2021.107700.
30. Deng Y, Lumley T. Multiple imputation through XGBoost. *J Comput Graph Stat.* 2024;33(2):352–63. doi:10.1080/10618600.2023.2252501.