International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Criteria for Measuring Intelligent Systems Quality in the Context of Contemporary International Standards

Mansoor N. Ali[1] , Motea M. Aljafare[2]

Submitted: 26/01/2024 **Revised**: 04/03/2024 **Accepted**: 12/03/2024

Abstract: Intelligent systems are more common in many facets of modern life, and whether they succeed or fail largely depends on how well they are made and how strictly they follow standards. First, standards must be established and evaluated for Intelligent systems. Organizations struggle to deploy itelligent systems efficiently because they lack explicit quality requirements. A crucial component of quality assurance is selecting the appropriate standards. quality metrics for intelligent systems will be defined. This research study examines the traits, creation, and development processes of Intelligent systems. It defines the term quality Intelligent Systems. It discusses the fundamental standards and procedures for determining the caliber of quality metrics for intelligent systems and the factors that have the most significant bearing on that caliber. This study addresses intelligent system quality, measures it and how it may be regulated, and illustrates the necessity for quality ISs as they become essential to everyday interactions and activities.

Keywords: *Intelligent Systems,* International *Standards, Measuring, Quality.*

1. Introduction

Intelligent systems (ISs) are becoming more prevalent in many areas of contemporary life, and their success or failure is primarily determined by how well-built and closely-regulated they are. The establishment and evaluation of standards are foundational requirements for the development of intelligent systems. Organizations find it difficult to implement ISs effectively due to the absence of clear quality standards. Choosing the appropriate standards is an essential part of quality assurance. This study will elucidate the concept of quality metrics. In order to tackle significant and reasonably complicated problems and produce consistent and dependable solutions over time, ISs offer a standardized methodological approach. Intelligence, as outlined in various dictionaries, denotes the capacity to comprehend, grasp, and derive insights from experiences, along with capabilities for learning and information retention.

AI-powered software systems are known as ISs. They differ from conventional commercial off-the-shelf software with decision assistance, like accounting information systems or enterprise resource planning software, in that they have at least two characteristics geared toward the end user. Notably, intelligent systems allow decision-making with cognitive capacities comparable to or even exceeding those of humans for specific tasks **[1].** Furthermore, these systems pick up knowledge from datasets comprising real-world observations and occurrences.

[1]*Faculty of Computing and IT, Sana'a University - Sana'a, Yemen*
ORCID ID : 0009-0007-1465-58
[2] *Faculty of Computing and IT, Sana'a University &Faculty of Computing and IT, University of Science and Technology-Sana'a, Yemen*
ORCID ID : 0000-0001-8137-5565

Through different stages in the model-building process, the biases or prejudices observed in the real-world settings permeate the systems **[2]**. Transparency, accountability, and fairness are essential components of ISs, particularly as these technologies proliferate in our daily lives graphical representations and written explanations are key tools in conveying complex concepts **[3]**. Software quality for information systems has been measured using a variety of models, including the ISO 25010, FURPS, Dromey, McCall, and Boehm models. Every model was created using a distinct principle or idea as its foundation **[4]**. The initial step in assessing software quality is to define the quality requirements model that will guide the specification, design, and implementation of the assessment processes. A method is then used to indicate the evaluation activities **[5]**.

Total Quality Management (TQM) heavily relies on data measurement and analysis. Metrics are used to assess the quality of methods, instruments, and products. They also account for effort and errors in quality performance. Additionally, they enable the development team to maintain effective development procedures **[6]**. Engineers seek quantitative performance measures to gauge the level of intelligence exhibited by a system **[7]**. The growth of the sector and the competitiveness of products on the global market are linked to the promotion and enhancement of artificial intelligence standards **[8]**. The applications and products of intelligent systems, which have human-like intelligence and even self-awareness due to their embedded intelligent algorithms or programs, have rapidly developed and are now incorporated into every facet of human existence **[9]**.

The ability to analyze data, understand the relationships between events or objects, carry out meaningful activities, and adapt learned knowledge to shifting circumstances are some of the prerequisites for a system to be considered intelligent [10]. An intelligent system of this type has to have the following characteristics: If it is robust, mobile and distributed, self-correcting, self-organizing, fault-tolerant, and adaptive, it should not fail (or at least fail gracefully) and be safe to use. Awareness of context and scenarios must be able to sense individuals.

2. Literature Review

2.1 Description of Standards

The International Organization for Standardization (ISO), the International Electrotechnical Commission (IEC), and the Institute of Electrical and Electronics Engineers (IEEE) have published a multitude of standards for systems engineering. A large number of the standards are mutually referenced **[11]**. These standards encompass the formulation of guidelines, norms, and procedures that regulate the development, application, and upkeep of intelligent systems. Through fostering transparency, responsibility, and collaboration, these guidelines aim to advance the ethical, safe, and reliable application of intelligent systems. The primary functions of information technology (IT) systems are covered by a number of standards, including ITIL, CMMI, and ISO 9001. Despite having unique advantages, none of them can meet all of the requirements for an IT system on its own **[12]**.

2.2 Quality measurement

Quality cannot be determined by a single metric. It demands the definition of characteristics and terms that can be used to define and evaluate quality standards. ISO/IEC 25010 delineates two models—one for product quality and another for quality in use. Regarding the quality of intelligent systems, the circumstances are comparable to those pertaining to quality here. Every developer has a unique understanding of the quality of intelligent systems, and they employ a different set of metrics (properties, factors, indicators) to measure that quality in practice **[13]**.

2.3 Intelligent Systems

The subject of ISs is difficult and contentious. A system's memory, learning, adaptability, adaptiveness, temporal dynamics, reasoning, and ability to handle vague and imprecise data can all be considered indicators of its intelligence from a computational perspective **[12]**.

2.4 Intelligent Systems Standards

Standards make it possible for different ISs to collaborate and communicate with one another without the requirement for specialist translation or integration. Standards promote quality by outlining exact requirements for dependability, performance, and safety. These components work together to accomplish common goals. Standards aid in defining best practices for ISs development, design, and implementation. Manufacturers and developers of ISs can ensure that their products fulfill defined quality criteria by adhering to these standards, which help to foster confidence and trust in the systems.

3. Elated Work

Talk about a few of the most recent international standards. When possible, provide a concise summary of each standard's contents that echoes its introduction. These ISs Compatibility and the IS International Standards Survey: -

3.1 SQuaRE

An assessment framework designed for evaluating software product quality is known as SQuaRE (Software Product Quality Requirements and Evaluation). It encompasses several key characteristics:

3.1.1 Functional appropriateness: This criterion evaluates the extent to which intelligent systems meet the defined functional requirements.

3.1.2 Maintainability: It measures how quickly intelligent systems may be modified, updated, or enhanced, focusing on the system's ability to support updates, bug fixes, and further changes.

3.1.3 Effectiveness: This characteristic assesses the ability of intelligent systems to properly and thoroughly assist users in achieving their goals while using the application in a certain environment.

3.1.4 Context Coverage: It refers to the degree to which intelligent systems are compatible within the particular context or setting in which they are designed to function.

3.1.5 The SQuaRE standard's aforementioned properties lead to the conclusion that it has some potential for use with intelligent systems.

3.2 DIN SPEC 92001-2 :

Is a German Institute of Standardization (DIN) standard that gives guidance for the creation and operation of trustworthy AI systems. It focuses on the ethical and legal elements of AI systems and strives to promote responsible AI use. The standard is significant to intelligent systems because it gives rules for developing and deploying these systems in an ethical, transparent, and accountable manner. It provides recommendations for ensuring that AI systems respect human autonomy, are clear and understandable, and comply with legal and ethical criteria

3.3 IEEE (ECPAIS –7010™ -2020 - P7014™) :

The IEEE principles include a wide range of intelligent system-related problems, such as data protection, transparency, accountability, justice, and human oversight. They also offer advice on how to design intelligent systems that are inclusive and accessible to all users, regardless of socioeconomic background, color, gender, or other characteristics.

3.4 ISO 25010 :

It serves as a standard for evaluating system product quality, providing a comprehensive framework for the assessment of system products. It covers various aspects of intelligent systems quality, such as functionality, reliability, performance, and usability, which are relevant to intelligent systems development**.** This standard defines two quality models, which describe desired quality characteristics of a system. The two models described in this standard are the “quality in use” model and the “product quality” model [10], which is a specification that lists a number of qualities for software systems and products. It contains:

3.4.1 Functionality: describes how well a system complies with requirements and carries out its intended functions.

3.4.2 Suitability: which system is appropriate for a certain use case or goal.

3.4.3 Reliability: It refers to the ability of system to perform its intended functions consistently and accurately over a specified period and under specified conditions.

3.4.4 Usability: It refers to how simple and effective a system's user interface is.

3.4.5 Compatibility: It is the system's capacity to cooperate and function as intended with other systems, hardware, software, or components

3.4.6 Performance: It shows the system's ability to achieve specified levels of response time, throughput, resource utilization, and other relevant metrics under specific conditions.

3.4.7 Efficiency: It is the capacity of a software system or product to achieve its goals with the least amount of resources used.

3.5 ISO/IEC 30141: This standard, titled "Artificial Intelligence -- Quality Evaluation of AI Systems," is currently under development. It aims to provide guidelines for evaluating the quality of AI systems, addressing aspects such as transparency, accountability, and robustness.

3.6 IEEE P7006:

This IEEE standard project focuses on defining ethical considerations for the design and deployment of autonomous and intelligent systems. It aims to promote transparency, accountability, and the responsible use of intelligent systems.

3.7 NIST SP 800-63B:

Although not specific to intelligent systems, the National Institute of Standards and Technology's (NIST) publication, SP 800-63B, provides guidelines for digital identity authentication. It specifically addresses the use of biometrics, including facial recognition systems and other biometric technologies used in intelligent systems.

Table 1. Survey of international standards and their compatibility with Intelligent systems

Modern Standards	***Field Standards***	***Standards Characteristics***		***Compatibility***
SQuaRE	Specifying, measuring and evaluating AI system quality	• Functional Suitability • Maintainability	• Effectiveness • Context coverage	√
DIN SPEC 92001-2	AI lifecycle process and quality	• Functionality • Performance	• Robustness • Comprehensibility	√
IEEE (ECPAIS – 7010™ -2020 - P7014™)	Include a wide range of intelligent system-related problems	• Protection • Accountability • Human Oversight	• Justice Transparency	√
ISO 25010	describes a collection of quality attributes and sub advantages that can be used to evaluate software products, especially those that incorporate intelligent systems.	• Functional • Suitability • Reliability • Usability • Compatibility	• Performance • Efficiency • Portability • Security • Maintainability	Same think
ISO/IEC 30141	The standard aims to provide guidance on evaluating the quality of AI systems throughout their lifecycle.	• Transparency • Fairness • Accountability	• Safety • robustness.	√
IEEE P7006	The standard focuses on addressing privacy, security, and ethical considerations associated with AI	• Privacy • Security • Ethics	• Accountability • Consent • Data Minimization	√

	systems that interact with personal data.			
NIST SP 800-63B	Assuring trustworthy and secure authentication methods in a variety of settings, including those that involve intelligent systems.	• Identity Proofing • Authentication • Assurance Levels	• Biometric Considerations • Lifecycle Management	√

4 Criteria For Intelligent Systems Quality

The quality of intelligent systems is a critical factor that can impact their effectiveness, safety, and overall success.

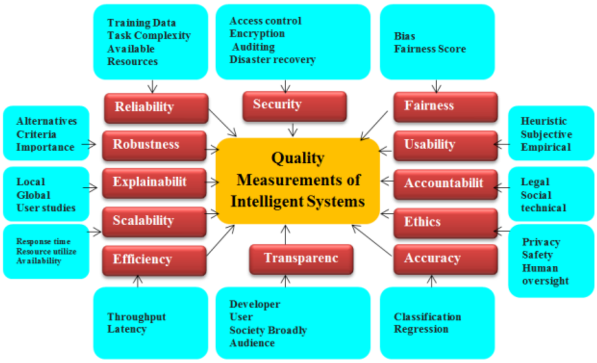

Fig 1. The Criteria for Quality Measurements of Intelligent Systems

Here are some important criteria for evaluating the quality of intelligent systems:

4.1 Reliability:

Intelligent systems should be reliable and consistent in their performance, even under varying conditions, measuring the reliability of intelligent systems is necessary to ensure that they are operating as planned. Selecting the strategy that best meets your needs can be made easier by being aware of the numerous dependability measurement approaches. Reliability is closely related to robustness and resilience, but the focus is on the time dimension. There are three key elements in the definition of reliability, including: failure, time, and environment. The failure events of an IS system can be mostly related to software errors, in addition to the failure of hardware **[14].** The reliability of an intelligent system can be affected by a number of factors, including:

4.1.3.1 The quality of the training data: If the data used to train the system is not accurate, it may be difficult for it to adapt to new circumstances.

4.1.3.2 Task complexity: The more complicated the task, the harder it is for the system to complete it correctly.

4.1.3.3 The available computational resources: A system may not be able to process the data rapidly enough to reach correct conclusions if it lacks sufficient computational resources.

Intelligent systems must be reliable and consistent in their performance, even under different conditions. There are some of the methods that are commonly used to measure the reliability of intelligent systems:

4.1.2.1 Mean time to failure (MTTF): This measures how long an intelligent system typically lasts before failing.

4.1.2.2 Mean time to repair (MTTR): This measures the typical time required to fix an intelligent system once it malfunctions.

4.1.2.3 Availability: This refers to the proportion of time that an intelligent system is accessible for use.

4.1.2.4 Reliability growth modeling is a method for projecting an intelligent system's reliability through time.

Table 2. Reliability Characteristics of *IS*

Variable	*Average*	*Composite Reliability (ρ)*	*Cornbrash Alpha (α)*
MTTF	0.795	0.886	0.742
MTTR	0.783	0.915	0.861
Availability	0.720	0.928	0.903
Growth modeling	0.893	0.962	0.940

Recurrent events data is utilized for the purpose of making predictions regarding reliability. In the context of AI reliability, the failure time can be the time to an incident that leads to a system failure caused by the AI systems. Such incident can arise from either hardware or software sources [14]. J.A.K. Suykens et al. presented a methodology in their article titled "A Comparative Study of Machine Learning Algorithms for Credit Scoring," published in IEEE Transactions on Neural Networks and Learning Systems in 2020, wherein they employed a scoring process to evaluate the reliability of intelligent systems [15].

Intelligent systems frequently employ the F1 score as a reliability metric, particularly when there is an imbalance between the dataset's classes. The F1 score is derived from the harmonic mean of recall and precision, two additional reliability metrics. The ratio of true positives (TP) to all positive predictions (TP + false positives, or FP) is known as precision. It calculates the percentage of optimistic forecasts that come true. The ratio of true positives to the total number of real positives, on the other hand, is calculated as follows: TP + FN. It calculates the percentage of real positives that the system accurately identifies.

The following formula can be used to gauge an intelligent system's dependability:

$$F1 = Reliability = 2 * \frac{precision * recall}{precision + recall}$$
(1)

For instance, consider a dataset consisting of 100 patients, with 90 being negative and 10 being positive. An intelligent system is trained on this dataset and used to predict the class labels of a new set of 50 patients. Let's assume that the system correctly predicts 8 out of the 10 positive patients but also misclassifies 5 negative patients as positive. In this scenario, the values of TP, FP, FN, and TN are as follows:

TP = 8 , FP = 5 , FN = 2 , TN = 35 (since the system correctly predicted 35 negative patients). Using these values, we can compute the precision and recall as follows:

Precision = TP / (TP + FP) = 8 / (8 + 5) = 0.615 ,

Recall = TP / (TP + FN) = 8 / (8 + 2) = 0.8

Subsequently, the F1 score can be computed using the formula:

Reliability = 2 * (0.615 * 0.8) / (0.615 + 0.8) = 0.696

Hence, the system achieves an F1 score of 0.696, which serves as a measure of the overall reliability of the system in accurately identifying positive patients.

4.2 Robustness: Intelligent systems should be able to handle unexpected situations and inputs gracefully, without crashing or producing incorrect results. Furthermore, the system should be able to function effectively even in the presence of unexpected inputs or changes to the operating environment, and intelligent systems should be robust to errors and unexpected inputs. This means that they should be able to continue to function properly even when presented with data that is incomplete, corrupted, or otherwise incorrect. There are two general approaches to robust AI: robust against model errors and robust against unmolded phenomena **[**16], making sure intelligent systems are functioning as intended requires measuring their resilience. You may be confident that you are utilizing the best way for your purposes by studying the many methods for assessing robustness. The following are some of the methods that are frequently used to gauge the robustness of intelligent systems:

The term "robustness" describes an intelligent system's ability to perform well in any operational environment, even with unexpected inputs or variables, without crashing or delivering inaccurate outputs. deciding on the alternatives, establishing the criteria, determining the relative weights of each criterion, and assessing each alternative's criterion. There are four components to the problem:

- Determine the R alternatives
- Set criteria C
- The relative importance (weights) of each criterion rt
- The criterion values for each alternative V Explainability

It can be expressed in the following formula:-

$$\text{Robustness} = (\mathrm{Ri} * \mathrm{C})\mathrm{rt} / \mathrm{V} \qquad (2)$$

4.3 Explainability: Human-understandable explanations of the thought and decision-making processes should be available from intelligent systems. Users should be able to comprehend how the system operates and why it generates the results that it does. In order to make intelligent systems visible and accountable, it is crucial to assess their explainability. By comprehending the various explainability measurement techniques, to determine how explainable intelligent systems are, a variety of methodologies are widely used:

- Local interpretability: It is a method for figuring out how a system decides in a particular situation.
- Global interpretability: This method helps us comprehend how a system decides in general.
- User studies: This method involves asking users to rank the system's explainability.

The explainability of an intelligent system can be affected by a number of factors, including:

- The system's type: Some systems are easier to understand than others. For instance, rule-based systems are frequently easier to understand than neural networks.
- The task that is expected to complete: Some tasks are easier to explain than others. Simple rules, for instance, are often easier to convey than sophisticated reasoning.
- The user's requirements: Different users require various levels of explanation. people who are unfamiliar with AI, for instance, could require more explanation than people who are.

In a broader context, explainability encompasses interpretability, which involves conveying the functioning of machine learning systems to users, as well as completeness **[18]**. A comprehensive review outlines five general desiderata for effective explanations of intelligent systems, contributing valuable insights to recent advancements in this field.

Table 3 General Desiderata for Useful Explanations of *ISs*

Code	*Description*
D1:	Fidelity the explanation must be a reasonable representation of what the system actually does.
D2:	Understandability involves multiple usability factors including terminology, user competencies, levels of

abstraction and interactivity.

D3: Sufficiency should be able to explain function and terminology and be detailed enough to justify decision.

D4: Low Construction overhead The explanation should not dominate the cost of designing AI.

D5: Efficiency the explanation system should not slow down the AI significantly.

Returning to the list of desiderata, several recent papers have aimed at framing the discourse of interpretability. They argue that interpretability lacks a well-defined concept and goes on to discuss multiple dimensions of interpretability and formulates a set of desiderata

4.4 Usability: Intelligent systems should be designed with a focus on usability and user experience. Usability, within the context of intelligent systems, refers to how well users can use the system to accomplish their objectives. Usability, a crucial component of the whole user experience, encompasses fundamental aspects such as simplicity, adaptability, effectiveness, contentment, and accessibility. Usability is recognized as a quality attribute and is defined by five key components [19]:

4.4.1 Learnability: This refers to the ease with which first-time users can manage to exercise all basic functionalities of the design.

4.4.2 Efficiency: It denotes the speed at which users can carry out their tasks once they are accustomed to the design.

4.4.3 Memorability: This pertains to the ease with which users can resume their former skills of site usage.

4.4.4 Errors: It encompasses the frequency, severity, and ease of recovery from user-made errors.

4.4.5 Satisfaction: This encompasses the enjoyment of using the design. Several techniques are available to analyze usability. It is important to note that the usefulness of an intelligent system can be influenced by various factors, including:

- User interface design: The usability of an intelligent system can be significantly affected by the design of its user interface.
- User experience: The overall perception and sentiment of users towards an intelligent system can impact its usefulness.
- Model accuracy: Users are more likely to trust and utilize a system that produces accurate results.

4.5 Scalability: refers to the capacity of intelligent systems to effectively handle large volumes of data and users, without sacrificing performance or accuracy. Scalability in intelligent systems is typically measured by their ability to efficiently handle increasing amounts of data, users, or tasks. To evaluate scalability, it's important to test the intelligent system under realistic conditions, simulating the expected workload and usage patterns. This can be done through load testing and performance testing, which involve simulating high levels of traffic or activity to see how the system performs under stress.

There are a few key metrics commonly used to assess the scalability of intelligent systems, including:

- Response time: This measures the time it takes for an intelligent system to respond to a user request or complete a task. As the system scales up, the response time should remain relatively constant or increase only moderately.
- Throughput: This measures the number of requests or tasks that an intelligent system can handle per unit of time. As the system scales up, the throughput should increase proportionally.
- Resource utilization: This measures how efficiently the intelligent system uses its available resources, such as CPU, memory, and storage. As the system scales up, it should be able to make more efficient use of its resources to handle the increased workload.
- Availability: This measures the percentage of time that the intelligent system is available and responsive to user requests. As the system scales up, it should maintain a high level of availability to ensure that users can always access it when needed

4.6 Efficiency: pertains to the ability of intelligent systems to use resources judiciously. This means that they should be able to produce results quickly and without using excessive amounts of memory or processing power. Efficiency in intelligent systems is typically measured by how effectively the system can accomplish its intended tasks while minimizing resource usage, such as CPU time, memory, or energy consumption.

To evaluate efficiency, it is crucial to consider these metrics within the context of the specific use case and application of the intelligent system. For example, a system that is highly accurate but slow in processing requests may not be efficient for real-time applications, while a system that is fast but inaccurate may not be efficient for data analysis tasks.

There are several key metrics that are commonly used to evaluate the efficiency of intelligent systems, including:

- Accuracy: This measures the correctness of the system's output or predictions, and is a key indicator of efficiency in systems that perform classification, prediction, or other types of data analysis.

- Throughput: This measures the rate at which the system can process requests or tasks, and is a key indicator of efficiency in systems that handle large volumes of data or requests.

- Resource utilization: This measures how effectively the system uses its available resources, such as CPU, memory, and storage. Efficient systems can accomplish their tasks with minimal resource usage, allowing them to scale effectively and reduce costs.

- Latency: This measures the time it takes for the system to process a request or complete a task, and is a key indicator of efficiency in systems that require real-time or near-real-time performance.

4.7 Security: The system should be designed and implemented with appropriate security measures to prevent unauthorized access, data breaches, and other security risks. Intelligent systems should be secure from unauthorized access or manipulation. This is especially important for systems that are used to control critical infrastructure or that contain sensitive data. Furthermore, the system's security characteristics, such as the reliability of traditional security detection, the effectiveness of security response mechanisms, and the precision of the security strategy, should be taken into consideration **[20].**

The security threats of intelligent systems can be limited to : Sneak Attacks(Se1), Probe or Scan (Se2), Automated Eavesdropping (Se3), Automated password attacks (Se4), spoofing (Se5), denial-of-service attacks (Se6), malware (Se7), physical infrastructure attacks (Se8), human error (Se9), and social engineering (Se10). Assuming that Q represents the number of requests made to the intelligent system, the calculation of the system's expectations can be expressed as follows:

$$\sum_{i=1}^{10} Se(i)/Q \quad (3)$$

4.8 Fairness: The system should be designed and implemented in a way that avoids bias and discrimination, and provides equal treatment to all users regardless of their race, gender, or other characteristics. Intelligent systems should be fair. This means that they should not discriminate against any particular group of people or individuals. Fairness can be broadly defined as the state of being impartial towards every individual and groups involved. However, fairness can be perceived differently by different people and contexts **[21]**.

It is crucial to standardize the bias measurement on a linear scale so that a uniform scale can be used to assess fairness and enable the comparison of different AI systems. Therefore, we introduce Bias Index for each protected attribute and Fairness Score for the overall system as the standard benchmarks for measuring fairness. Fairness Score is defined for the AI system as follows [21]:

$$\text{FS} = 1- \sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(M\,ij-Mj)2}{mn}} \quad (4)$$

where, i : number of the protected attribute , j : number of the fairness metric , n : total number of fairness metrics used , m : total number of protected attributes considered in the AI system , Mij : value of the jth fairness metric for the ith protected attribute , Mj0 : ideal value of the jth fairness metric , i:e: 0 for difference metrics and 1 for ratio metrics.

Transparency: To ensure that users comprehend how the system arrived at its results, its decision-making procedures should be clear and explicable. Users should be able to see through intelligent systems. This means that the system's operation and the reasons behind the outputs it generates should be clear to users. For some tasks, intelligent systems allow decision-making with cognitive capacities comparable to or even greater than those of humans [22]. However, a growing body of design-based research on explainable intelligent systems contends that the opaque nature of deep learning algorithms makes users reluctant to use the systems, which reduces their effectiveness[22]. Interest in providing more efficient system training, more dependability, and enhanced usability has increased as a result of transparency and accountability [22]. Humans' tendency to mistrust AI predictions is a common barrier to the adoption of assistive AI systems. For this reason, the field studying artificial intelligence has been concentrating on making AI decisions more understandable by offering justifications **[21]**.

The principle of transparency mandates the ease of accessibility and comprehension of all information and communications pertaining to the processing of personal data, together with the use of unambiguous language and various forms of transparency:

- To the developer.
- To the user
- To the community at large
- To provide an expert.
- To facilitate monitoring, testing and the public.

- **Accountability**: It is imperative that intelligent systems exhibit accountability, which refers to the ability to hold system developers or owners responsible for the actions performed by the system. This entails the obligation to provide information about actions taken, offer explanations or justifications for those actions, and take subsequent actions, including punishment or rectification [21]. In the context of intelligent systems, accountability pertains to the responsibility of designers, operators, and users to assume ownership of the system's decisions and outcomes. This includes accountability for whatever damage the system may have caused as well as accountability for making sure the system functions in a safe, just, and efficient manner. Intelligent systems are made to decide or act based on data and algorithms, which can have a big impact on people, businesses, and society as a whole. For instance, a mistaken autonomous vehicle could result in a serious collision.

- Organizational accountability: This refers to the obligation placed on businesses that create or use intelligent systems to make sure that the right procedures and frameworks are in place to control the system's risks. Among the essential forms of accountability are:

- Legal accountability: This refers to the responsibility individuals or organizations bear under the law for the activities and outcomes of intelligent systems.

- Social accountability: In order to ensure that intelligent systems are created and used in a way that is consistent with the needs and values of society as a whole, people and organizations have a responsibility known as social accountability.

- Technical accountability: This entails the obligation placed on people or organizations to make sure that intelligent systems are developed and deployed in a safe and secure manner.

- **Ethics**: Intelligent systems should be designed and used in an ethical and responsible manner, with attention paid to issues such as bias, privacy, and fairness. These system should adhere to ethical principles and values, and should not be used to support activities that are illegal, harmful, or unethical. In April 2016, the IEEE Standards Association launched a global initiative on the Ethics of Autonomous and Intelligent Systems. The significance of this initiative cannot be overstated; coming from a professional body with the standing and reach of the IEEE Standards Association, it marks a watershed in the emergence of ethical standards. and it is a radical step **[15]**. Some significant ethical guidelines for intelligent systems include :

- **Privacy**: Intelligent systems should be developed to safeguard the confidentiality of users' personal information. As a result, developers must take action to make sure that data is gathered and handled in accordance with recognized privacy standards and laws.

- **Safety:** Emphasizing the utmost priority on safety during the development of intelligent systems, efforts should be made to minimize the risk of harm to individuals, property, or the environment. This is especially critical for systems such as autonomous vehicles or medical equipment that have the potential to cause harm.

- **Human oversight**: Intelligent systems should be built with human monitoring and intervention in mind, especially when making decisions or taking actions that could have a big impact on people or society as a whole

- **Accuracy:** Intelligent systems should continually deliver correct outcomes with a low error rate , discuss several techniques for assessing the precision of intelligent systems, as well as the difficulties in doing so. They also shed light on the variables that can influence the accuracy of intelligent systems, like the standard of the training data and the difficulty of the task at hand.

To make sure that intelligent systems are functioning as intended, it is crucial to assess their accuracy. Understanding the various approaches to measuring precision can help you choose the one that is best for your requirements . The way accuracy is measured relies on the type of intelligent system in question and the use case for which it is designed. Here are a few typical techniques for gauging the precision of intelligent systems:

- Classification accuracy: It is a measurement of how often an intelligent system classifies something correctly. Usually, this is determined by comparing the system's predictions to a list of predetermined results or labels.

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \times 100\% \quad (5)$$

- Regression accuracy: The degree to which an intelligent system's anticipated values match the actual values is measured by regression accuracy. Usually, this is calculated by contrasting the system's predictions with a predetermined set of values.

$$\text{R-squared} = 1 - \frac{\text{Sum of squared residuals}}{\text{Total sum of squares}} \quad (6)$$

5 Factors that Affect Quality Measures for Intelligent System

As intelligent systems get more complicated and powerful and are utilized in a variety of applications, it is crucial to make sure that they are created and used in a way that is morally upright, open, and responsible. In order to do this, it may be necessary to evaluate the effectiveness and impact of intelligent systems using a variety of quality metrics. Table (4) lists the variables and elements that influence smart system quality measurements and that regulate them according to a certain set of rules when they change.

Table 4 Factors Effect the Measurements of Intelligent systems

Variables	***Items***	***Measurement Method***
Reliability	Training Data Task Complexity Available Resources	- Calculate the quality of the data used to train the system - Determine the complexity. - The computational resources
Robustness	Alternatives Criteria Importance	- Identify hostile attacks - Make distributional transformations of data - Determine data quality - Know the complexity of the model - Clarify the system architecture
Explainability	Interpretability (Local - Global - User Studies	- Decide on a local situation - Decide on a position in general. - Explain the system to users
Usability	Heuristic Subjective Empirical	- Good user interface design - users experience - Adjust the accuracy of the model
Scalability	Response Time Resource Utilize	- measurement of reaction time - The extent of resource use
Efficiency	Throughput Latency	- Assessing the veracity of the outputs of the system - Monitoring system performance - Resource efficiency assessment
Security	Access Control Encryption Auditing Disaster Recovery	- Limit access and exercise control - Data protection - What is vulnerability management? - assemble threat intelligence - Planning for disaster recovery
Fairness	Bias Fairness Score	- Bias testing - Fairness metrics - Adversarial testing
Transparency	Developer User Audience	- Explainability methods: - Model interpretability: - Auditing involves examining
Accountability	Legal - Social Technical Organizational	- Governance frameworks - Performance metrics - Regulatory compliance
Ethics	Privacy Safety Human Oversight	- protection of privacy - Putting safety first - human interference and oversight
Accuracy	Classification Regression	- classification accuracy . - measurement Regression precision

6 Conclusion

In conclusion, the significance of adhering to standards and implementing quality control measures cannot be overstated for the success of ISs in various aspects of contemporary life. Organizations are unable to use ISs successfully since there are no clear quality criteria for them. However, there are a number of worldwide standards for software and systems engineering as well as developing standards for machine learning and artificial intelligence that can be applied to the construction of intelligent systems. Recent standards, such as the IEEE ECPAIS recommendations and the IEEE 7010-2020 standard, provide a framework for creating ethically responsible AI and autonomous systems that align with human values. Researchers, policymakers, and organizations involved in the development and deployment of intelligent systems can benefit from the proposed framework for assessing the quality of intelligent systems within the context of modern international standards. The proposed framework can be a useful tool for organizations developing and deploying intelligent systems, as well as researchers and policymakers interested in ensuring that intelligent systems are developed and deployed ethically, transparently, and in alignment with human values. The adoption of the proposed framework enables organizations to design and develop their intelligent systems in a manner that upholds responsibility and accountability, ultimately contributing to their success across various areas of modern life. Notably, intelligent systems standards emphasize the utilization of models and algorithms that are interpretable and transparent, particularly when addressing concerns such as prejudice, discrimination, privacy, security, and ethical responsibility. These standards also promote the establishment of consistent and interoperable data exchange between multiple systems. Furthermore, intelligent systems standards encourage the development and implementation of new and improved technologies, techniques, and best practices, fostering innovation and continuous improvement in the field. By adhering to these standards, organizations can navigate the ethical and societal implications associated with intelligent systems, while promoting openness, interpretability, and the advancement of intelligent system capabilities.

References

[1] Wanner, Jonas, et al. The effect of transparency and trust on intelligent system acceptance: Evidence from a user-based study. Electronic Markets, 2022, 1-24.

[2] Agarwal, Avinash; Agarwal, Harsh; Agarwal, Nihaarika. Fairness Score and process

standardization: framework for fairness certification in artificial intelligence systems. AI and Ethics, 2023, 3.1: 267-279.

[3] Smith-Renner, Alison Marie, et al. TExSS: Transparency and explanations in intelligent systems. In: 26th International Conference on Intelligent User Interfaces-Companion. 2021. p. 24-25.

[4] Abidin, Wan Yusran Naim Wan Zainal, and Zulkefli Mansor. "The Criteria for Software Quality in Information System: Rasch Analysis." Editorial Preface From the Desk of Managing Editor… 10.9 (2019).

[5] Pérez, L. S. V., Tornés, A. F. G., Riverón, E. M. F., Hernández, A. M. S., & Galeana, N. I. P. (2018). Software Quality Methodology to Train Engineers as Evaluators of Information Systems Development Tools. Universal Journal of Educational Research, 6(12), 2942-2951.

[6] Aouhassi, S., & Hanoune, M. (2018). Information System Quality: Managers Perspective. International Journal Of Advanced Computer Science And Applications, 9(8), 493-502.

[7] GAO, Rong; Tsoukalas, L. Performance metrics for intelligent systems: An engineering perspective. Nist Special Publication SP, 2002, 5-10.

[8] Russia, Evgeniy Bryndin. Standardization of Artificial Intelligence for the Development and Use of Intelligent Systems. Advances in Wireless Communications and Networks, 2020, 6.1: 1.

[9] YU, Rong. Security Framework of Artificial Intelligence System. In: Journal of Physics: Conference Series. IOP Publishing, 2021. p. 012011.

[10] Dahiya, Pawan, et al. Intelligent systems: Features, challenges, techniques, applications & future scopes. Intelligent Systems & Mobile Adhoc Networks, 2007.

[11] Schneider, Florian; Berenbach, Brian. A literature survey on international standards for systems requirements engineering. Procedia Computer Science, 2013, 16: 796-805.

[12] Dahar, Hind; Roudies, Ounsa. Mapping of Quality Standards. IJITEE, 2019, 8: 2406-2412.

[13] Kopyltsov, A. V. (2020, April). Selection of metrics in software quality evaluation. In Journal of Physics: Conference Series (Vol. 1515, No. 3, p. 032018). IOP Publishing.

[14] Hansen, Lars Kai; Rieger, Laura. Interpretability in intelligent systems–a new concept?. Explainable AI: Interpreting, Explaining and Visualizing Deep Learning, 2019, 41-49.

[15] Barmer, Hollen, et al. Robust and Secure AI. 2021.

[16] Hamon, Ronan; Junklewitz, Henrik; Sanchez, Ignacio. Robustness and explainability of artificial intelligence. Publications Office of the European Union, 2020, 207.

[17] Dingli, Alexiei; CASSAR, Sarah. An intelligent framework for website usability. Advances in Human-Computer Interaction, 2014, 2014: 5-5.

[18] Knauer, T., Nikiforow, N., & Wagener, S. (2020). Determinants of information system quality and data quality in management accounting. Journal of Management Control, 31(1), 97-121

[19] RUDAS, Imre J.; FODOR, János. Intelligent systems. International Journal of Computers, Communications & Control, 2008, 3.3: 132-138.

[20] Li, Z., Wu, S., Zhou, H., Zou, S., & Dong, T. (2019, August). Analytic model and assessment framework for data quality evaluation in state grid. In Journal of Physics: Conference Series (Vol. 1302, No. 2, p. 022083). IOP Publishing.

[21] Hong, Yili, et al. Statistical perspectives on reliability of artificial intelligence systems. Quality Engineering, 2023, 35.1: 56-78.

[22] A Deep Learning Framework for Predicting Tumor Proliferation Score of Breast Histopathological Images" by Yujing Wang et al., published in the IEEE Journal of Biomedical and Health Informatics in 2020.