ISO Form 4
NEW WORK ITEM PROPOSAL (NP)

Circulation date: 2024-10-29 **Closing date for voting:** 2025-01-27	**Reference number:** ISO/IEC NP TS 25570 ISO/IEC JTC 1/SC 42 **N 1954**
Proposer ISO/IEC JTC 1/SC 42	
Secretariat ANSI	

A proposal for a new work item within the scope of an existing committee shall be submitted to the secretariat of that committee.
A proposal for a new project committee shall be submitted to the Central Secretariat, which will process the proposal in accordance with ISO/IEC Directives, Part 1, Clause 2.3.
Guidelines for proposing and justifying new work items or new fields of technical activity (Project Committee) are given in ISO/IEC Directives, Part 1, Annex C.
IMPORTANT NOTE: Proposals without adequate justification and supporting information risk rejection or referral to the originator.

☒ The proposer confirms that this proposal has been drafted in compliance with Annex C of ISO/IEC Directives, Part 1.

PROPOSAL

(to be completed by the proposer, following discussion with committee leadership if appropriate)

TITLE

English title:

Information Technology -- Artificial Intelligence -- Reliability assessment of AI systems

French title:

(In the case of an amendment, revision or a new part of an existing document, show the reference number and current title)

SCOPE

This document provides methods and mechanisms to assess the reliability of an AI system. It describes the metrics of reliability and the procedure for reliability assessment from a statistical perspective.

PURPOSE AND JUSTIFICATION

With the wide spread roll out of AI systems in every aspect of human life, it is important to assess the reliability of AI systems before and during each real-world deployment. This is especially important for AI systems that affect various aspects of human life, such as health-care, robotic surgery, autonomous vehicles, automated senior citizen monitoring and care, etc. Reliability assessment is important because failure events of an AI system can lead to business loss, information loss, human injury and other safety issues. With large-scale deployment of an AI system, assessing reliability can also help system administrators to have a level of confidence in the functioning of that AI system before roll-out and during its life-cycle.

Reliability assessment focuses on estimating how well the AI system can perform its designed functionality without failure, for the intended period of time, under given conditions for operational profiles.

Reliability models can give a predictive measure that the system would function at a level of performance for a period of time in a given environment. High reliability can help consumers and users be confident of the AI system against potential failures during run-time of the system. This is important for all AI systems, especially the ones that can have a direct impact on human life and safety.

Reliability is estimated by analysing all failure data of the system, using statistical modelling techniques leading to building an estimate of the potential future failure prediction in various scenarios.

Reliability can be viewed as related to quality and testing of AI systems, but is a very different aspect. While testing determines whether the AI system's output matches the expected output, reliability estimates the confidence in the system to function without failures for a specified period of time after development or deployment of the AI system based on the failure logs of the system. Quality, as described in ISO/IEC 25059:2023, has reliability as one of the characteristics of quality model of AI system. The Quality model evaluates the AI system based on functional and non-functional characteristics, and is a broader term that can include various aspects such as functional adaptability, transparency, intervenability, societal and ethical risk mitigation etc.

It is important that for AI systems we assess its reliability for failure free operation for a specified period of time under stated conditions. By doing this one can have a certain level of confidence that the system would function well as required and not create failures and faults during its run. With automated systems coming into every aspect of human life it is important that we document processes through which this reliability-level can be measured and published for AI systems. For autonomous AI systems, core decision making AI systems and critical care AI systems it is suggested that reliability of such systems be measured and reported before the AI system goes live, and also during its life in different operating environments. Reliability assessment does not determine whether an AI output is factually correct, fair, safe, secure, ethical or robust. In reliability assessment the focus is on estimating the failure-free operation of the AI system for a given time period in a specific context.

Many national and regional bodies have listed reliability of AI systems as an essential and important aspect for responsible and trustworthy AI deployment.

1. India - the NITI Aayog Responsible AI approach document (Part 1 – Principles for Responsible AI, February 2021) under the Principle of Safety and Reliability, requires the AI system to be reliable for deployment.
2. US – in the NIST Artificial Intelligence Risk Management Framework (AI RMF 1.0, Jan 2023), Valid and Reliable is listed as a necessary condition of trustworthiness of AI systems.
3. Japan - the Guidelines on Assessment of AI Reliability in the Field of Plant Safety (Second edition, March 2021), emphasizes the need of a systematic methodology to assess the reliability of AI systems especially for critical sectors.

4. Republic of Korea - the National Strategy for Artificial Intelligence (October 2019), states that establishing a quality management system that verifies reliability of AI systems is required for the prevention of AI dysfunction.
5. China - the Ethical Norms for New Generation Artificial Intelligence (September 2021) in Article 12 lists reliability of the AI system as an important aspect.
6. UK - the roadmap to an effective AI assurance ecosystem (December 2021), on page 3 lists reliable as one of the concerns of AI systems to work as intended.
7. Australia - the Artificial Intelligence Ethics Framework describes reliability of AI systems as one of the ethics principles to ensure that AI systems reliably operate in accordance with their intended purpose throughout their lifecycle.
8. Canada – the Principles for responsible, trustworthy and privacy-protective generative AI technologies (Office of the Privacy Commissioner of Canada, Dec 7, 2023) mandates that organizations using generative AI should evaluate the reliability of generative AI tool for the intended purpose.
9. The EU's High Level Expert Group on Artificial Intelligence published a document (dated 8 April 2019) on Ethics Guidelines for Trustworthy AI that lists reliability of AI systems as a critical aspect.
10. In the EU, Artificial Intelligence Act (AIA-Final-Draft-21-January-2024.pdf) in Article 7(2), point (ga) the mechanisms for reliability of AI system is mentioned as an important aspect.
Following ISO/IEC documents mention the reliability of AI systems:
In ISO/IEC 22989:2022(E), in clause 3.5.9 reliability is defined as 'property of consistent intended behaviour and results'. Reliability is also mentioned in 3.5.16 as a Note 2 to entry on trustworthiness that mentions reliability as a characteristic for trustworthiness. As per clause 5.15.3 reliability of an AI system refers to the ability that enables it to provide required prediction, recommendation and decision consistently correctly during its operation stage.
ISO/IEC 25059:2023, in clause 5.1, AI System Product Quality, reliability is one of the characteristics of quality model of AI system. Maturity, Availability, Fault tolerance, Recoverability and Robustness are the sub-characteristics of Reliability. Some of the sub-characteristics, such as robustness of neural networks, has been addressed in other SC42 projects and those projects would be appropriately referenced in this proposed project.
ISO/IEC 25023:2016, Clause 8.6 describes reliability measures with Maturity measures comprising of MTBF, Failure rate etc. that can be used for reliability assessment.
ISO/IEC 23894:2023, Reliability is listed as a Risk source in Annex-B.
ISO/IEC TR 5469:2024, reliability is mentioned in Clause, 8.2, 9.3.6 and 10.3 as an essential part of the functional safety of the AI system.
ISO/IEC TS 5723:2022 Reliability is defined as "ability of an item to perform as required, without failure, for a given time interval, under given conditions"
ISO/IEC TR 24028:2020(E), clause 10 states that Reliability of AI systems play a very vital role in developing trust in AI systems.
ISO/IEC DIS 42005:2024- Clause-5.8.2.6, Table C.1 and Section F mentions reliability as one of the Impact assessment aspect.
ISO/IEC FDIS 5392:2023- mentions reliability as one of the concerns of KE Stakeholders (7.2)
Thus, it is very important to develop a document related to reliability assessment of AI systems. While software reliability measures are well established, reliability measures for AI systems need to be described. This proposed technical specification will describe methods and measures for assessing the reliability of AI systems so that it is measured and reported. This can be done at any time after the testing of the AI system or while it is being deployed or while it is in real-world use.
The project will develop the reliability characteristic as listed in ISO/IEC 25059:2023, and will describe how the reliability of an AI system can be assessed based on some of the maturity measures of reliability in ISO/IEC 25023:2016,

Sustainable Development Goals (SDGs)

Goal 9: Industry, Innovation, and Infrastructure

Preparatory work

☐ A draft is attached ☒ An outline is attached ☐ An existing document serving as the initial basis is attached

The proposer is prepared to undertake the preparatory workrequired:

☒ Yes ☐ No

If a draft is attached to this proposal:

Please select from one of the following options:

☒ The draft document can be registered at Preparatory stage (WD – stage 20.00)
☐ The draft document can be registered at Committee stage (CD – stage 30.00)
☐ The draft document can be registered at enquiry stage (DIS – stage 40.00)

If the attached document is copyrighted or includes copyrighted content:

☐ The proposer confirms that copyright permission has been granted for ISO to use this content in compliance with the ISO/IEC Directives, Part 1 (see also the Declaration on copyright).

Is this proposal for an ISO management System Standard (MSS)?

☐ Yes ☒ No

Note: If yes, this proposal must have an accompanying justification study. Please see the Consolidated Supplement to the ISO/IEC Directives, Part 1, Annex SL or Annex JG

Indication of the preferred type to be developed

☐ International Standard ☒ Technical Specification

☐ Publicly Available Specification *

* While a formal NP ballot is not required to start developing a PAS (no eForm04), the NP form may provide useful information for the committee P-members to consider when deciding to initiate a Publicly Available Specification.

Proposed Standard Development Track (SDT – to be discussed by the proposer with the committee manager or ISO/CS)

☐ 18 months ☐ 24 months ☒ 36 months

Draft project plan (as discussed with committee leadership)

Proposed date for first meeting: 2025-02-03

Dates for key milestones: Circulation of 1st Working Draft (if any) to experts: 2025-02-28
Committee Draft consultation (if any): 2025-09-30
DIS submission*:
Publication*: 2027-05-15

* Target Dates for DIS submission and Publication should be set a few weeks ahead of the limit dates automatically determined when selecting the SDT.

NOTE: ISO/Meetings and ISO/Projects allow you to register and continuously update the meeting dates and project target dates during the development of the project.

Known patented items (see ISO/IEC Directives, Part 1 for important guidance)

☐ Yes ☒ No

If "Yes", provide full information as annex

Co-ordination of work: To the best of your knowledge, has this or a similar proposal been submitted to another standards development organization?

☐ Yes ☒ No

If "Yes", please specify which one(s):

Listing of relevant documents (such as standards and regulations) at international, regional and national level

Identification and description of relevant affected stakeholder categories (Please see ISO CONNECT)

	Benefits/Impacts/Examples
Industry and commerce - large industry	Stakeholders will be able to evaluate the reliability of an AI system.
Industry and commerce - SMEs	Stakeholders will be able to evaluate the reliability of an AI system
Government	Users of the standards in setting policies and regulations as the authority. Can set up policies and guidelines on AI system and its expected reliability.
Consumers	Users will be able to know about the reliability of an AI system before and during use.
Labour	Provide framework for understanding the reliability of AI system
Academic and research bodies	This document could be the basis for further research on the methods for reliability of AI systems.
Standards application businesses	Some consumer advocacy, civil societies and organizations will find this standard helpful in understanding the stakeholders' roles and responsibilities in the reliability of AI systems
Non-governmental organizations	
Other (please specify)	

Liaisons:	**Joint/parallel work:**
A listing of relevant external international organizations or internal parties (other ISO and/or IEC committees) to be engaged as liaisons in the development of the deliverable.	**Possible joint/parallel work with:** ☐ IEC (please specify committee ID) ☐ CEN (please specify committee ID) ☐ Other (please specify)

A listing of relevant countries which are not already P-members of the committee.

Note: The Committee Manager shall distribute this NP to the ISO members of the countries listed above to ask if they wish to participate in this work

Proposed Project Leader (name and e-mail address)	**Name of the Proposer** (include contact information)
C. Anantaram c.anantaram@gmail.com	Heather Benko hbenko@ansi.org

This proposal will be developed by:

☒ An existing Working Group: ISO/IEC JTC 1/SC 42/WG 3 Trustworthiness

☐ A new Working Group:

(Note: establishment of a new Working Group requires approval by the parent committee)

☐ The TC/SC directly

☐ To be determined:

Supplementary information relating to the proposal

☒ This proposal relates to a new ISO document

☐ This proposal relates to the adoption as an active project of an item currently registered as a Preliminary Work Item

☐ This proposal relates to the re-establishment of a cancelled project as an active project

Other:

Maintenance agencies (MA) and registration authorities (RA)

☐ This proposal requires the designation of a maintenance agency. If so, please identify the potential candidate:

☐ This proposal requires the designation of a registration authority. If so, please identify the potential candidate:

NOTE: Selection and appointment of the MA or RA are subject to the procedure outlined in ISO/IEC Directives, Part 1, Annex G and Annex H.

☒ Annex(es) are included with this proposal (provide details)

Additional information/question(s)