Enhancing Robustness and Security of Edge AI Systems for Safety-Critical Applications

WP5 – Impact Maximization: Dissemination, Exploitation, and Certification Roadmap

D5.4: Review of EU and international standards

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101168067. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Document Information

GRANT AGREEMENT NUMBER	101168067	**ACRONYM**	GuardAI
FULL TITLE	Enhancing Robustness and Security of Edge AI Systems for Safety-Critical Applications		
START DATE	1st October 2024	**DURATION**	36 months
PROJECT URL	https://www.kios.ucy.ac.cy/guardai/		
DELIVERABLE	D5.4: Review of EU and international standards		
WORK PACKAGE	WP5 – Impact Maximization: Dissemination, Exploitation, and Certification Roadmap		
DATE OF DELIVERY	**CONTRACTUAL** 02/2026 (M17)	**ACTUAL**	02/2026
TYPE	R — Document, report	**DISSEMINATION LEVEL**	PU - Public
LEAD BENEFICIARY	CYS		
RESPONSIBLE AUTHORS	Marios Mavroyiannos, Stefanos Gurov, Joseph Karis, Konstantinos Zaou, Stylianos Petrou		
CONTRIBUTIONS (FROM)	Marios Mavroyiannos (CYS), Stefanos Gurov (CYS), Joseph Karis (CYS), Konstantinos Zaou (CYS), Stylianos Petrou (CYS), Christos Kyrkou (UCY-KIOS CoE), Antonis Savva (UCY-KIOS CoE)		
ABSTRACT	This report describes the methodology used in the GuardAI project to identify, screen, and select European and international standards relevant to its technical objectives. A structured, multi-phase approach was applied: mapping relevant standardisation bodies and collecting 886 standards, followed by classification by publication status and relevance-based screening. Through title- and scope-based analysis and collaborative review, a focused subset of standards with high technical and strategic relevance was identified and documented in the K4AI Knowledge Hub. In addition, two standards (ISO/IEC 24029-1 and ISO/IEC 24029-2) were selected for in-depth analysis and presented at the project's third standardisation seminar, as they are directly applicable to assessing the robustness of neural networks in safety-critical AI systems.		

Document History

VERSION	ISSUE DATE	STAGE	DESCRIPTION	CONTRIBUTOR
V 0.1	19/11/2025	Draft	ToC	Marios Mavroyiannos (CYS), Stefanos Gurov (CYS), Joseph Karis (CYS), Konstantinos Zaou (CYS), Stylianos Petrou (CYS),
V0.2	19/12/2025	Draft	Initial Contributions	Marios Mavroyiannos (CYS), Stefanos Gurov (CYS), Joseph Karis (CYS), Konstantinos Zaou (CYS), Stylianos Petrou (CYS),
V0.3	19/01/2026	Draft	Final Contributions	Christos Kyrkou (UCY-KIOS CoE), Antonis Savva (UCY-KIOS CoE)
V0.4	06/02/2026	Draft	Comments from internal reviewers	Marios Mavroyiannos (CYS), Stefanos Gurov (CYS), Joseph Karis (CYS), Konstantinos Zaou (CYS), Stylianos Petrou (CYS), Erion-Vasilis Pikoulis (ATHENA), Christos Anagnostopoulos (ATHENA), Nikos Piperigkos (ATHENA), Christos Mavrokefalidis (ATHENA), Alexandros Gkillas (ATHENA), Aris Lalos (ATHENA)
V1.0	26/02/2026	Final	Final edits	Marios Mavroyiannos (CYS), Stefanos Gurov (CYS), Joseph Karis (CYS), Konstantinos Zaou (CYS), Stylianos Petrou (CYS), Christos Kyrkou (UCY-KIOS CoE), Antonis Savva (UCY-KIOS CoE)

Disclaimer

Copyright message

Table of acronyms

Acronym	Meaning
CEN	European Committee for Standardization
CENELEC	European Committee for Electrotechnical Standardization
ETSI	European Telecommunications Standards Institute
JTC	Joint Technical Committee
K4AI	Knowledge for Artificial Intelligence
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
ITU-T	International Telecommunication Union - Telecommunication Standardization Sector
MTS	Methods for Testing and Specification
SAI	Securing Artificial Intelligence
SDO	Standards Development Organization
SC	Sub-Committee
TC	Technical Committee

Table of Contents

Executive summary

This report documents the methodology used in the GuardAI project to systematically identify, evaluate, and select standards relevant to its objectives. The activity aimed to ensure that GuardAI research and innovation activities are informed by the state of the art in international and European standardisation, while also supporting regulatory alignment and future exploitation.

A structured, multi-phase methodology was applied, beginning with a broad scoping of relevant standardisation bodies and technical committees, followed by progressive filtering based on relevance to the project's technical focus, and concluding with a collaborative evaluation and prioritization process.

The outcomes of this work are reflected in the list of public and draft standards published in the K4AI Knowledge Hub. Additionally, an in-depth analysis of two selected standards was performed and presented during the project's 3rd plenary meeting, a seminar titled "Assessing Neural Network Robustness: ISO/IEC 24029 Standards (part 1 & 2)".

1. Introduction

GuardAI is a project focused on enhancing the robustness and security of Edge AI systems for safety-critical applications. Given the regulatory and technical sensitivity of AI systems deployed in such contexts, alignment with recognised standards is critical to ensuring trustworthiness, robustness, and compliance.

The purpose of this report is to establish a transparent and repeatable methodology for identifying and selecting standards directly relevant to the technical scope and objectives of GuardAI. Rather than relying on ad-hoc or informal selection, a structured process was adopted to ensure traceability, completeness, and consistency across project outputs.

The screening of standards included two discrete approaches. Selecting standards that could support the implementation of the project tasks and deliverables, and standards that the project could possibly influence in shaping by proposing suggestions/modifications stemming from the implementation work.

2. Standards identification and selection methodology

2.1. Introduction

The screening methodology for selecting relevant standards to the projects objectives and outputs included three (3) discrete stages:

- **Stage 1**: **Initial search and scoping** – Identification of relevant Technical Committees under ISO, IEC, CEN, CENELEC, ETSI and list of all standards drafted by them.
- **Stage 2**: **Filtering and initial screening** – Screen standards based on standard title and further narrow down based on standards scope ensuring technical relevance.
- **Stage 3**: **Collaborative Evaluation and Final Selection**. Review with consortium technical experts and perform a final screening with actionable items.

In addition to assessing the technical relevance of standards, the applied methodology explicitly considered the regulatory and strategic context of AI standardisation in Europe.

The objective was to ensure that the standards examined were not only aligned with the technical needs of the GuardAI project but also relevant to ongoing and forthcoming regulatory and policy developments. Emphasis was placed on standards addressing trustworthiness, robustness, and cybersecurity of AI systems, as these areas are central to both the project's scope and the evolving European AI regulatory landscape.

This approach ensured that the selected standards supported immediate project implementation needs while also facilitating future compliance, interoperability, and the uptake of project results within the broader European AI ecosystem.

2.2. Stage 1: Initial search and scoping

Stage 1 of the screening focused on identifying and scoping standardisation activities relevant to AI. This phase aimed to establish a comprehensive, structured baseline of standards against which subsequent screening and refinement could be conducted.

2.2.1. Identification of Technical Committees

The stage initiated from the mapping of relevant Technical Committees (TCs) and Subcommittees (SCs). The mapping was conducted across major international and European Standard Developing Organisations (SDOs), namely ISO, IEC, CEN, CENELEC, and ETSI. Priority was given to committees with active work programmes in AI technologies, information security, cybersecurity, system robustness, testing methodologies, and regulatory support standards.

Based on this mapping, the following technical committees were identified as the primary sources of relevant standards for GuardAI:

- ***ISO/IEC JTC 1/SC 42 – Artificial Intelligence***

  - ***ISO/IEC JTC 1*** Sub-committee 42 (SC 42) is responsible for standardization in the field of Artificial Intelligence. It serves as the central focus of JTC 1's AI program and provides guidance to other ISO and IEC committees developing AI applications.

- ***ISO/IEC JTC 1/SC 27 – Information Security, Cybersecurity and Privacy Protection***
  - ***ISO/IEC JTC 1*** Sub-committee 27 (SC 27) develops standards for the protection of information and ICT, covering security requirements, management systems, cryptographic mechanisms, and privacy aspects. It also engages in active liaison and collaboration with other bodies to ensure the proper application of its standards and technical reports.

- ***CEN/CLC JTC 21 – Artificial Intelligence***
  - The CEN and CENELEC Joint Technical Committee 21 (JTC 21) is dedicated to developing European standards for Artificial Intelligence (AI), including harmonized standards in support of the EU AI Act. Established on June 1, 2021, JTC 21 unites over 300 experts from more than 20 countries, working collaboratively through five specialized working groups.

- ***ETSI TC SAI – Securing Artificial Intelligence***
  - The ETSI Technical Committee Securing Artificial Intelligence (TC SAI) develops specifications to mitigate threats from and to AI systems. Its scope encompasses securing AI against attacks, mitigating malicious AI, leveraging AI to enhance security, and addressing societal security considerations. The committee focuses initially on Machine Learning while providing guidance on broader AI developments.

- ***ETSI TC MTS – Methods for Testing and Specification***
  - The ETSI Technical Committee Methods for Testing and Specification (TC MTS) creates standards for testing and specification languages. It provides frameworks and methodologies to support other ETSI committees, working closely with the Centre for Testing and Interoperability to develop foundational materials for ETSI and bodies like ITU-T Study Group 17. Its work has also been widely adopted by global organizations and the industry.

These committees collectively represent the core international and European standardisation activities relevant to the project's technical and regulatory scope.

2.2.2. Collection of Standards References

For each identified technical committee, *all available standard references* under its responsibility were manually listed into a single, structured dataset. This involved a thorough, manual review and compilation process of the committees' official publications and work programmes. This process **identified 886 standards** originating from the five committees listed above.

To enable systematic analysis, the collected standards were categorised according to their publication status, resulting in the following groups:

- **Published standards** (734)
- **Draft standards** (152)

Following this classification, the standards were further organised by title and thematic focus, allowing an initial screening for relevance to GuardAI objectives.

2.3. Stage 2: Filtering and initial screening

The second stage focused on narrowing the initially broad corpus of identified standards to a subset more relevant to GuardAI's objectives, based on the standards' titles and scopes.

To support this process, a structured table was developed to systematically capture key information for each standard, including the following:

- reference or standard number,
- title,
- short description or scope,
- publication status,
- harmonisation status,
- category,
- issuing body,
- relevant EU policies,
- legislation or mandates, and
- whether the standard could be listed in the K4AI Knowledge Hub.

As part of this stage, the standards were further organised into two distinct groups based on their publication status. The first group comprised published standards, from which 179 standards were identified as relevant following the initial filtering.

The second group included standards not yet published, encompassing draft standards, standards under approval, and other ongoing standardisation work items; it comprised 77 standards.

2.3.1. Title-Level Screening

An initial manual, expert-driven screening of standards titles was conducted to assess their potential relevance to the project's objectives. The keywords used to perform the screening were:

- AI robustness,
- trustworthiness,
- cybersecurity of AI systems, and
- reliability of safety-critical systems.

The screening resulted in the following standards:

- 152 Published standards
- 37 Draft standards

It is noted that, at this stage of screening the draft standards, the title relevance was further considered, along with the standardisation process status and timing, to include draft standards that align with the project timelines. In this way, the project can contribute to these draft standards while respecting standardisation procedures.

2.3.2. Scope Verification

For standards passing the title-level screening, the official **scope descriptions** were examined. Where necessary, **key sections** of the standards were reviewed to verify alignment with the project's technical focus.

The screening resulted in the following standards:

- 28 Published standards
- 19 Draft standards

The official scope descriptions of these standards were reviewed to assess their relevance to key GuardAI themes, including

- AI robustness,
- trustworthiness,
- cybersecurity,
- testing methodologies, and
- assessment of AI systems.

This scope-based screening excluded standards that were not applicable to the project's domain and use cases.

2.4. Stage 3: Collaborative Evaluation and Final Selection

2.4.1. Overview

The third stage of the methodology is built directly on the outcomes of the second stage. The stage focused on further consolidating and prioritising the **28 published and 19 draft standards** retained after the second screening phase.

Through collaborative evaluation, these standards were assessed for technical relevance and strategic importance to GuardAI, yielding a focused, actionable shortlist from which the standards selected for in-depth analysis were identified.

The final screening resulted in the following standards:

- **28 Published standards for K4AI listing**
- **2 Published standards for consortium training**
- **4 Draft standards**

2.4.2. Review by Project Coordinator and final refinement

The shortlist of the standards resulting from the second stage was subsequently submitted to the project coordinator for review. This review focused on prioritisation, identification of overlaps, and assessment of the relative technical and strategic importance of each standard within the project.

Based on feedback and follow-up discussions with the standards body, the shortlist was further refined to identify standards that offer the greatest added value for in-depth analysis and direct use within GuardAI.

The process resulted in the publication of **28 standard references** in the K4AI (Knowledge Hub for AI).

Furthermore, **2 published standards** were selected as the most relevant for detailed examination, analysis, and presentation within the project, as part of the 3rd standardisation seminar.

Lastly, **4 draft standards** with the highest potential to have a substantial impact on the project were shortlisted from the 19 draft standards.

The detailed final list of identified standards is provided in the Appendix, which covers their classification, publication status, harmonization status, and legislative relevance. The appendix includes separate tables for published and draft standards, containing information that can be made publicly available.

3. Dissemination of Identified Standards

3.1. Introduction

The K4AI Knowledge Hub consolidates the outcomes of the project's standards identification and selection process, offering a structured overview of the international standards most relevant to AI robustness, cybersecurity, and system reliability. These standards—presented in dedicated tables for ISO/IEC, CEN, and ETSI—form the foundational reference base guiding GuardAI's alignment with globally recognised best practices.

Among them, two ISO/IEC standards (24029-1 and 24029-2) were selected for in-depth analysis due to their direct relevance to the project's focus on neural network robustness in safety-critical applications.

Additionally, the project adopts ISO/IEC 5259-1:2025 to ensure high-quality, compliant, and traceable data processes within Use Case 1. Together, these standards reinforce GuardAI's commitment to building trustworthy, certifiable, and resilient AI systems.

3.2. K4AI Knowledge Hub

The outcomes of the standards identification and selection methodology are consolidated and published within the **K4AI Knowledge Hub,** under AI Certification (https://www.kios.ucy.ac.cy/guardai/certification-and-evaluation/).

To ensure clarity and usability, the standards are listed in 3 discrete tables based on the issuing body, i.e., ISO/IEC, CEN, and ETSI.

This list includes the **28 published standards** identified during the final screening stage. These standards provide essential background, terminology, frameworks, and established practices relevant to AI, cybersecurity, and system robustness.

While not all published standards were analysed in depth, they form the broader reference pool supporting the project's understanding of the existing standardisation landscape.

3.3. Focus on the Two Selected Standards (Standardisation seminar 3)

The selection of **ISO/IEC 24029-1** and **ISO/IEC 24029-2** for in-depth analysis and presentation at the third project standardisation seminar was driven by their direct relevance to GuardAI's core technical objectives. Both standards address the assessment of neural network robustness, a critical property for AI systems deployed in safety-critical and security-sensitive environments, which are central to the project's scope.

Standard	Topic	Rationale for Selection
ISO/IEC 24029-1	Assessment of robustness of neural networks – Overview	Provides foundational concepts for AI robustness evaluation relevant to safety-critical systems
ISO/IEC 24029-2	Assessment of robustness of neural networks – Methodology	Defines concrete methodologies for evaluating robustness, directly supporting GuardAI objectives

ISO/IEC 24029-1 was selected as it provides a high-level overview and conceptual foundation for robustness assessment of neural networks. The standard establishes common terminology, key robustness concepts, and an overall framework to support a shared understanding of robustness in AI systems. This perspective is particularly important for the project, as it enables consistent interpretation of robustness requirements across different use cases, partners, and technical components within GuardAI.

ISO/IEC 24029-2 was selected as a complementary standard because it builds on the conceptual framework of Part 1 and provides concrete methodologies for robustness assessment. It defines practical approaches, testing considerations, and evaluation principles that can be directly applied when analysing neural network behaviour under perturbations, adversarial conditions, or unexpected inputs. This methodological focus aligns closely with GuardAI's technical work on evaluating and improving the robustness and security of edge AI models.

Taken together, the two standards form a coherent, complementary pair: Part 1 establishes the conceptual basis, while Part 2 operationalizes these concepts into actionable assessment methods. Their combined presentation during Webinar 3 enabled the consortium to demonstrate how existing standardisation work can be leveraged at both the theoretical and practical levels, supporting the project's research activities and reinforcing alignment with recognised international standards in AI robustness.

These two standards directly support GuardAI by providing a two-layered standardization basis for evaluating the robustness of neural networks deployed in autonomous vehicles, drones, and 5G edge infrastructure. ISO/IEC 24029-1 establishes the foundational terminology, concepts, and assessment framework that

enables GuardAI to consistently and standards-aligned define robustness properties under a unified taxonomy and conceptual model. ISO/IEC 24029-2 complements this by specifying testing techniques and evaluation principles for quantifying neural network resilience against input perturbations and adversarial attacks (mapping directly to GuardAI's objective of developing resilient edge AI algorithms). Together, the two parts supply the structure that GuardAI requires to formulate standardized evaluation criteria and certification frameworks, ensuring that robustness claims for its target systems are defined, measured, and validated against an internationally recognized reference baseline.

3.4. Adoption of Standards for Data Quality Assurance

Beyond the standards mentioned, the project implements **ISO/IEC 5259-1:2025 (Artificial Intelligence - Data quality for analytics and machine learning)** as the framework to validate and operationalize the data collection strategy outlined in *Deliverable D1.3: Reporting on data collection, requirements, and handling* for *Use Case 1: Robust Surveillance and Monitoring with AI-Enabled Drones*.

Use Case 1 leader translated the standard into enforceable data-quality metrics and processing controls across every data lifecycle stage. It defined explicit image-stability criteria: frames at a 1/12 ratio were screened, excluding those with significant motion blur. Only stable, high-fidelity images were entered into the annotation pipeline, thereby operationalizing ISO/IEC 5259-1's requirement that training data meet quality criteria relevant to the target task.

To ensure annotation consistency, the project used CVAT as the only annotation tool, utilizing its predictive features to reduce variability. Manual quality checks on each batch verified label accuracy and spatial precision before exporting to YOLO and Pascal VOC formats. These protocols ensure repeatable, auditable data processes, providing traceable evidence of quality control.

The standard's data governance and privacy provisions require de-identification controls: all human faces and vehicle license plates captured during aerial surveillance were blurred before processing or publication, ensuring GDPR compliance and ISO/IEC 5259-1's privacy measures proportional to content sensitivity.

This adherence to ISO/IEC 5259-1:2025 ensures that the dataset meets safety-critical machine-learning standards. It can be confidently used for training and deploying reliable AI models in demanding aerial surveillance scenarios.

4. Conclusion

This report presents a structured, transparent methodology for identifying, screening, and selecting European and international standards relevant to the GuardAI project's objectives. Through a phased approach, ranging from initial scoping and classification to targeted filtering and collaborative evaluation, the methodology reduced a broad standardisation landscape to a focused, actionable set of standards. The outcomes support both the project's immediate technical needs and its strategic alignment with ongoing and emerging AI standardisation activities, while ensuring traceability, consistency, and relevance throughout the process.

Appendix

Table 1 – Published standards

A/A	Reference /Standard Number	Title	Short description/ Scope	Publication Status	Harmonization status	Category	Issuing Body	Relevant EU policy/ legislation/ Mandates
1.	CEN/CLC ISO/IEC/TR 24027:2023	Information technology - Artificial intelligence (AI) - Bias in AI systems and AI aided decision making (ISO/IEC TR 24027:2021)	This document addresses bias in AI systems, particularly in AI-aided decision-making. Measurement techniques and methods for assessing bias are described to address bias-related vulnerabilities. All AI system lifecycle phases are in scope, including but not limited to data collection, training, continual learning, design, testing, evaluation, and use.	Published	NO	Governance and quality of datasets used to build AI systems	CEN/CLC/JTC 21	Non-harmonized Standard / Optional Use for AI Act
2.	CEN/CLC ISO/IEC/TR 24029-1:2023	Artificial Intelligence (AI) - Assessment of the robustness of neural networks - Part 1: Overview (ISO/IEC TR 24029-1:2021)	This document provides background on existing methods for assessing the robustness of neural networks.	Published	NO	Robustness specifications for AI systems	CEN/CLC/JTC 21	Non-harmonized Standard / Optional Use for AI Act
3.	CEN/CLC ISO/IEC/TS 12791:2024	Information technology - Artificial intelligence - Treatment of unwanted bias in classification and regression machine learning tasks (ISO/IEC TS 12791:2024)	This document describes how to mitigate bias in AI systems that use machine learning for classification and regression tasks. It provides mitigation techniques applicable throughout the AI system lifecycle to address unwanted bias. This document is applicable to all types and sizes of organizations.	Published	YES	Governance and quality of datasets used to build AI systems	CEN/CLC/JTC 21	AI Act

A/A	Reference /Standard Number	Title	Short description/ Scope	Publication Status	Harmonization status	Category	Issuing Body	Relevant EU policy/ legislation/ Mandates
4.	CEN/CLC/TR 18115:2024	Data governance and quality for AI within the European context	This document provides an overview of AI-related standards, with a focus on data and data life cycles, for organizations, agencies, enterprises, developers, universities, researchers, focus groups, users, and other stakeholders navigating this era of digital transformation. It describes links among the many international standards and regulations published or under development, with the aim of promoting a common language, a greater culture of quality, and providing an information framework. It addresses the following areas: - data governance; - data quality; - elements for data, data sets properties to provide unbiased evaluation and information for testing.	Published		Governance and quality of datasets used to build AI systems	CEN/CLC/JTC 21	AI Act
5.	EN ISO/IEC 23053:2023	Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML) (ISO/IEC 23053:2022)	This document establishes an Artificial Intelligence (AI) and Machine Learning (ML) framework for describing a generic AI system built with ML technologies. The framework describes the system components and their functions in the AI ecosystem. This document is applicable to all types and sizes of organizations, including public and private companies, government entities, and not-for-profit organizations, that are implementing or using AI systems.	Published	YES	Supporting standards (terminology)	CEN/CLC/JTC 21	AI Act
6.	EN ISO/IEC 23894:2024	Information technology - Artificial intelligence - Guidance on risk management (ISO/IEC 23894:2023)	This document provides guidance on how organizations that develop, produce, deploy, or use products, systems, and services that use artificial intelligence (AI) can manage AI-related risks. The guidance also aims to help organizations integrate risk management into their AI-related activities and functions. It describes processes for the effective implementation and integration of AI risk management. This guidance can be customized to any organization and its context.	Published	YES	Risk management for AI systems	CEN/CLC/JTC 21	AI Act

A/A	Reference /Standard Number	Title	Short description/ Scope	Publication Status	Harmonization status	Category	Issuing Body	Relevant EU policy/ legislation/ Mandates
7.	EN ISO/IEC 5259-1:2025	Artificial intelligence - Data quality for analytics and machine learning (ML) - Part 1: Overview, terminology, and examples (ISO/IEC 5259-1:2024)	This document provides a framework for understanding and associating the individual documents of the ISO/IEC 5259 series and serves as the foundation for conceptual understanding of data quality for analytics and machine learning. It also discusses associated technologies and examples (e.g., use cases and usage scenarios).	Published	YES	Governance and quality of datasets used to build AI systems	CEN/CLC/JTC 21	AI Act
8.	EN ISO/IEC 5259-2:2025	Artificial intelligence - Data quality for analytics and machine learning (ML) - Part 2: Data quality measures (ISO/IEC 5259-2:2024)	This document specifies a data quality model, data quality measures, and guidance on reporting data quality in the context of analytics and machine learning (ML). This document applies to all types of organizations seeking to achieve their data quality objectives.	Published	YES	Governance and quality of datasets used to build AI systems	CEN/CLC/JTC 21	AI Act
9.	EN ISO/IEC 5259-3:2025	Artificial intelligence - Data quality for analytics and machine learning (ML) - Part 3: Data quality management requirements and guidelines (ISO/IEC 5259-3:2024)	This document specifies requirements and provides guidance for establishing, implementing, maintaining, and continually improving the quality of data used in analytics and machine learning. This document does not define a detailed process, methods, or metrics. Rather, it defines the requirements and guidance for a quality management process along with a reference process and methods that can be tailored to meet the requirements in this document. The requirements and recommendations set out in this document are generic and intended to apply to all organizations, regardless of type, size, or nature.	Published	YES	Governance and quality of datasets used to build AI systems	CEN/CLC/JTC 21	AI Act
10.	EN ISO/IEC 5259-4:2025	Artificial intelligence - Data quality for analytics and machine learning (ML) - Part 4: Data quality process framework (ISO/IEC 5259-4:2024)	This document establishes general common organizational approaches, regardless of the type, size, or nature of the applying organization, to ensure data quality for training and evaluation in analytics and machine learning (ML). It includes guidance on the data quality process for: – supervised ML with regard to the labelling of data used for training ML systems, including common	Published	YES	Governance and quality of datasets used to build AI systems	CEN/CLC/JTC 21	AI Act

A/A	Reference /Standard Number	Title	Short description/ Scope	Publication Status	Harmonization status	Category	Issuing Body	Relevant EU policy/ legislation/ Mandates
			organizational approaches for training data labelling; – unsupervised ML; – semi-supervised ML; – reinforcement learning; – analytics. This document applies to training and evaluating data from different sources, including data acquisition, data composition, data preparation, data labelling, evaluation, and data use. This document does not define specific services, platforms, or tools.					
11.	EN ISO/IEC 8183:2024	Information technology - Artificial intelligence - Data life cycle framework (ISO/IEC 8183:2023)	This document defines the stages and identifies associated actions for data processing throughout the artificial intelligence (AI) system life cycle, including acquisition, creation, development, deployment, maintenance, and decommissioning. This document does not define specific services, platforms, or tools. This document applies to all organizations, regardless of type, size, or nature, that use data in the development and deployment of AI systems.	Published	YES	Governance and quality of datasets used to build AI systems	CEN/CLC/JTC 21	AI Act
12.	ETSI TR 103 910 V1.1.1 (2025-02)	Methods for Testing and Specification (MTS); AI Testing; Test Methodology and Test Specification for ML-based Systems	This document describes test types, test items, quality criteria, and testing methodologies for ML-based systems, with an emphasis on supervised, unsupervised, and reinforcement learning. The present document outlines how these testing practices can be effectively integrated into the life cycle of typical ML-based systems. This document applies to all types of organizations involved in any lifecycle stage of developing and operating ML-based systems, as well as to other stakeholder roles.	Published	NO	Security in AI Systems	ETSI TC MTS	Non-harmonized Standard / Optional Use for AI Act

A/A	Reference /Standard Number	Title	Short description/ Scope	Publication Status	Harmonization status	Category	Issuing Body	Relevant EU policy/ legislation/ Mandates
13.	ETSI TS 104 050 V1.1.1 (2025-03)	Securing Artificial Intelligence (SAI); AI Threat Ontology and definitions	The present document defines what an Artificial Intelligence (AI) threat is and defines how it can be distinguished from any non-AI threat. The AI threat model is presented as an ontology that provides a view of the relationships among actors representing threats, threat agents, assets, and related entities, and defines those terms. The ontology in this document builds on the base taxonomy of threats and threat agents described in ETSI TS 102 165-1 and addresses the overall problem statement for SAI presented in ETSI TR 104 221 [i.21] and the mitigation strategies described in ETSI TR 104 222 [i.22].	Published	NO	Security in AI Systems	ETSI TC SAI	Non-harmonized Standard / Optional Use for AI Act
14.	ETSI TR 104 221 V1.1.1 (2025-01)	Securing Artificial Intelligence (SAI); Problem Statement	This document describes the problem of securing AI-based systems and solutions, with a focus on machine learning, and the challenges related to confidentiality, integrity, and availability at each stage of the machine learning lifecycle. It also describes broader challenges in AI systems, including bias, ethics, and explainability. Several attack vectors are described, along with real-world use cases and attacks.	Published	NO	Security in AI Systems	ETSI TC SAI	Non-harmonized Standard / Optional Use for AI Act
15.	ETSI TR 104 048 V1.1.1 (2025-01)	Securing Artificial Intelligence (SAI); Data Supply Chain Security	This document addresses security issues arising from data supply chains in the development of Artificial Intelligence (AI) and Machine Learning (ML) systems. Data is a critical component in the development of AIML systems. Compromising data integrity has been shown to be a viable attack vector against such systems. This document summarizes the methods used to source data for AI training and reviews existing initiatives to develop data-sharing protocols. It then provides a gap analysis of these methods and initiatives to scope potential standards requirements to ensure the integrity and confidentiality of shared data, information, and feedback. This document primarily addresses data security rather than model security. It is recognized, however, that AI supply chains can be complex and that models can themselves be part of the supply chain, generating new data for subsequent training. Model security	Published	NO	Security in AI Systems	ETSI TC SAI	Non-harmonized Standard / Optional Use for AI Act

A/A	Reference /Standard Number	Title	Short description/ Scope	Publication Status	Harmonization status	Category	Issuing Body	Relevant EU policy/ legislation/ Mandates
			is therefore influenced by, and in turn influences, the security of the data supply chain. Mitigation and detection methods can be similar across data and models, with poisoning in one being detected by analyzing the other. The present document focuses on security; however, data integrity is not only a security issue. Techniques for assessing and understanding data quality for performance, transparency, or ethical purposes are applicable to security assurance as well. An adversary's aim can be to disrupt or degrade a model's functionality to achieve destructive effects. The adoption of security controls will likely improve performance and transparency, and vice versa. This document does not address data theft, which is a traditional cybersecurity problem. The focus instead is on data manipulation and its impact on AI/ML systems.					
16.	ETSI TR 104 222 V1.2.1 (2024-07)	Securing Artificial Intelligence; Mitigation Strategy Report	This document summarizes and analyzes existing and potential mitigation measures against threats to AI-based systems. The goal is to conduct a technical survey to mitigate threats introduced by adopting AI in systems. The technical survey shed light on available methods for securing AI-based systems by mitigating known and potential security threats. It also addresses security capabilities, challenges, and limitations when implementing AI-based mitigation in certain use cases.	Published	NO	Security in AI Systems	ETSI TC SAI	Non-harmonized Standard / Optional Use for AI Act

A/A	Reference /Standard Number	Title	Short description/ Scope	Publication Status	Harmonization status	Category	Issuing Body	Relevant EU policy/ legislation/ Mandates
17.	ETSI TR 104 066 V1.1.1 (2024-07)	Securing Artificial Intelligence; Security Testing of AI	This document identifies methods and techniques appropriate for security testing of ML-based components. Security testing of AI does not end at the component level. For traditional software testing, integration with other system components must also be tested. However, integration testing is not the focus of this document. The present document addresses: - security testing approaches for AI; - security test oracles for AI; - definition of test adequacy criteria for security testing of AI. Techniques of each of these topics should be applied together to a security test of an ML component. Security testing approaches generate test cases that are executed against the ML component. Security test oracles compute a test verdict indicating whether a test case has passed (no vulnerability detected) or failed (a vulnerability detected). Test adequacy criteria assess overall progress and can be used to specify a stop condition for security testing.	Published	NO	Security in AI Systems	ETSI TC SAI	Non-harmonized Standard / Optional Use for AI Act
18.	ETSI TR 104 062 V1.2.1 (2024-07)	Securing Artificial Intelligence; Automated Manipulation of Multimedia Identity Representations	This document covers AI-based techniques for automatically manipulating existing or generating synthetic identity data across media formats, such as audio, video, and text (deepfakes). This document describes various technical approaches and analyzes the threats posed by deepfakes across different attack scenarios. It then provides technical and organizational measures to mitigate these threats and discusses their effectiveness and limitations.	Published	NO	Security in AI Systems	ETSI TC SAI	Non-harmonized Standard / Optional Use for AI Act

A/A	Reference /Standard Number	Title	Short description/ Scope	Publication Status	Harmonization status	Category	Issuing Body	Relevant EU policy/ legislation/ Mandates
19.	ETSI TR 104 225 V1.1.1 (2024-04)	Securing Artificial Intelligence TC (SAI); Privacy aspects of AI/ML systems	This document identifies privacy as a component of AI security and defines measures to protect and preserve privacy in AI, covering both model safeguarding and data protection, as well as the role of privacy-sensitive data in AI solutions. It documents and addresses attacks and their associated remediations, where applicable, considering the existence of multiple trust levels across the data lifecycle. The attack mitigations investigated include non-AI-specific (traditional Security/privacy controls) and AI/ML-specific remedies, proactive remediations ("left of the boom"), and reactive responses to adversarial activity ("right of the boom").	Published	NO	Security in AI Systems	ETSI TC SAI	Non-harmonized Standard / Optional Use for AI Act
20.	ISO/IEC 18033-1:2021	Information security — Encryption algorithms — Part 1: General	This document is general in nature and provides definitions that apply in subsequent parts of the ISO/IEC 18033 series. It introduces the concept of encryption and outlines key aspects of its use and properties.	Published	NO	Security in AI Systems	ISO/IEC JTC 1/SC 27	Non-harmonized Standard / Optional Use for AI Act
21.	ISO/IEC 18033-6:2019	IT Security techniques — Encryption algorithms — Part 6: Homomorphic encryption	This document specifies the following mechanisms for homomorphic encryption. • Exponential ElGamal encryption; • Paillier encryption. For each mechanism, this document specifies the process for: • generating parameters and the keys of the involved entities; • encrypting data; • decrypting encrypted data; and • homomorphically operating on encrypted data. Annex A defines the object identifiers assigned to the mechanisms specified in this document. Annex B provides numerical examples.	Published	NO	Security in AI Systems	ISO/IEC JTC 1/SC 27	Non-harmonized Standard / Optional Use for AI Act

A/A	Reference /Standard Number	Title	Short description/ Scope	Publication Status	Harmonization status	Category	Issuing Body	Relevant EU policy/ legislation/ Mandates
22.	ISO/IEC 20546:2019	Information technology — Big data — Overview and vocabulary	This document provides a set of terms and definitions to improve communication and understanding in this area. It provides a terminological foundation for big data-related standards. This document provides a conceptual overview of the field of big data, its relationship to other technical areas and standards efforts, and concepts commonly associated with big data that are not new.	Published	NO	Supporting standards (terminology)	ISO/IEC JTC 1/SC 42	Non-harmonized Standard / Optional Use for AI Act
23.	ISO/IEC 24029-2:2023	Artificial intelligence (AI) — Assessment of the robustness of neural networks — Part 2: Methodology for the use of formal methods	This document presents a methodology for applying formal methods to assess the robustness properties of neural networks. The document focuses on how to select, apply, and manage formal methods to prove robustness properties.	Published	NO	Robustness specifications for AI systems	ISO/IEC JTC 1/SC 42	Non-harmonized Standard / Optional Use for AI Act
24.	ISO/IEC 42005:2025	Information technology — Artificial intelligence (AI) — AI system impact assessment	This document provides guidance for organizations performing artificial intelligence (AI) impact assessments for individuals and societies that can be affected by an AI system and its foreseeable applications. It includes considerations on how and when to perform such assessments, at which stages of the AI system life cycle, and guidance on AI system impact assessment documentation. Additionally, this guidance includes how this AI system impact assessment process can be integrated into an organization's AI risk management and AI management system. This document is intended for use by organizations developing, providing, or using AI systems. This document is applicable to any organization, regardless of size, type, or nature.	Published	NO	Governance and quality of datasets used to build AI systems	ISO/IEC JTC 1/SC 42	Non-harmonized Standard / Optional Use for AI Act

A/A	Reference /Standard Number	Title	Short description/ Scope	Publication Status	Harmonization status	Category	Issuing Body	Relevant EU policy/ legislation/ Mandates
25.	ISO/IEC 5338:2023	Information technology — Artificial intelligence — AI system life cycle processes	This document defines a set of processes and associated concepts for describing the life cycle of AI systems based on machine learning and heuristic systems. It is based on ISO/IEC/IEEE 15288 and ISO/IEC/IEEE 12207 with modifications and additions of AI-specific processes from ISO/IEC 22989 and ISO/IEC 23053. This document outlines processes for defining, controlling, managing, and executing the AI system across its life-cycle stages. These processes can also be used within an organization or a project when developing or acquiring AI systems. When an element of an AI system is traditional software or a traditional system, the software life cycle processes in ISO/IEC/IEEE 12207 and the system life cycle processes in ISO/IEC/IEEE 15288 can be used to implement that element.	Published	NO	Governance and quality of datasets used to build AI systems	ISO/IEC JTC 1/SC 42	Non-harmonized Standard / Optional Use for AI Act
26.	ISO/IEC TR 24028:2020	Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence	This document surveys topics related to trustworthiness in AI systems, including the following: • approaches to establish trust in AI systems through transparency, explainability, controllability, etc.; • engineering pitfalls and typical associated threats and risks to AI systems, along with possible mitigation techniques and methods; and • approaches to assess and achieve availability, resiliency, reliability, accuracy, safety, security, and privacy of AI systems. The specification of levels of trustworthiness for AI systems is out of the scope of this document.	Published	NO	Robustness specifications for AI systems	ISO/IEC JTC 1/SC 42	Non-harmonized Standard / Optional Use for AI Act

A/A	Reference /Standard Number	Title	Short description/ Scope	Publication Status	Harmonization status	Category	Issuing Body	Relevant EU policy/ legislation/ Mandates
27.	ISO/IEC TR 24368:2022	Information technology — Artificial intelligence — Overview of ethical and societal concerns	This document provides a high-level overview of AI ethical and societal concerns. In addition, this document: • provides information in relation to principles, processes, and methods in this area; • is intended for technologists, regulators, interest groups, and society at large; • is not intended to advocate for any specific set of values (value systems). This document provides an overview of International Standards that address AI-related ethical and societal concerns.	Published	NO	Transparency and information to the users of AI systems	ISO/IEC JTC 1/SC 42	Non harmonized Standard / Optional Use for AI Act
28.	ISO/IEC TR 5469:2024	Artificial intelligence — Functional safety and AI systems	This document describes the properties, related risk factors, available methods, and processes relating to: • use of AI inside a safety-related function to realize the functionality; • use of non-AI safety-related functions to ensure safety for an AI controlled equipment; • use of AI systems to design and develop safety-related functions.	Published	NO	Robustness specifications for AI systems	ISO/IEC JTC 1/SC 42	Non harmonized Standard / Optional Use for AI Act

Table 2 – Draft standards

Reference /Standard Number	Title	Short description/ Scope	Publication Status	Forecasted voting date	Harmonization status	Category	Issuing Body	Relevant EU policy/ legislation/ Mandates	Potential means of contribution
ISO/IEC AWI TS 22440-1	Artificial intelligence — Functional safety and AI systems — Part 1: Requirements	This document provides **requirements** and guidance on the terminology, properties, **risk factors**, processes, methods, techniques, and architectures relating to: - **use of AI technology within a safety-related function;** - use of safety-related functions based on conventional technology to ensure the safety of a system using AI technology; - use of AI technology to design, develop, and verify safety-related functions. This document includes general considerations on how security threats can affect the safety of an AI system. Unless otherwise specified, this document is applicable to all types of AI technologies. It includes specific details on machine learning.	New project registered in TC/SC work programme	TBD	No	Cybersecurity specifications for AI systems	ISO/IEC JTC 1/SC 42	No	Present Use Cases / Requirements / Risks (attack surfaces) that these systems may entail / Defence methods for each Use Case flow experiment.

Reference /Standard Number	Title	Short description/ Scope	Publication Status	Forecasted voting date	Harmonization status	Category	Issuing Body	Relevant EU policy/ legislation/ Mandates	Potential means of contribution
DTR/SAI-0014 (TR 104 197)	AI Functional Safety and Security	The intent of this work item is to identify the actions of AI developers to ensure that their products and systems remain safe and secure when deployed. This work item will report on the core measures and how they work alongside the core principles of SAI. The intent is to ensure systems are both secure and safe in the deployed context. This reinforces **security-by-design**. To also address how to give assurance of a secure and safe design	Committee draft (CD) registered - Work item adopted (2025-06-10)	TBD	No	Cybersecurity specifications for AI systems	ETSI TC SAI	No	Work done in the context of WP3: - standard profile format that includes adversarial robustness indicators (e.g., accuracy drop under attack). - methodology to assess defence mechanisms for AI models documenting the process, metrics, and computing requirements needed to produce profiling outputs (accuracy, fps, etc.) - estimation formulas to project metrics, given some input parameters, and provide recommendations based on the library of available defences and their properties.
prEN 18229(WI=JT021008)	AI trustworthiness framework	This document provides a framework for AI systems' trustworthiness that includes terminology, concepts, high-level horizontal requirements, guidance, and a method for contextualizing them for specific stakeholders, domains, or applications. The **high-level horizontal requirements address foundational aspects and characteristics of AI system trustworthiness**. This document is primarily intended for organizations **that place AI systems on the market or put them into service**.	Under Drafting	2026-09-30	YES	Cybersecurity specifications for AI systems	CEN/CLC/ JTC 21	YES	- Work done in the context of survey (D2.1) - Use Case leaders may be considered suppliers for the respective models trained.

Reference /Standard Number	Title	Short description/ Scope	Publication Status	Forecasted voting date	Harmonization status	Category	Issuing Body	Relevant EU policy/ legislation/ Mandates	Potential means of contribution
prEN XXX(WI=JT021025)	Artificial Intelligence – Evaluation methods for accurate computer vision systems	This document specifies the evaluation of computer vision systems, focusing on measuring the quality of a system's results to assess its functional suitability. It provides definitions of evaluation methods for those systems, along with guidance on selecting, implementing, and interpreting them. This document covers quantitative metrics and other evaluation methods. It includes requirements for implementing the described metrics and additional requirements for the technical resources used in the evaluation process.	Under Drafting	2026-09-02	YES	Accuracy specifications for AI systems	CEN/CLC/ JTC 21	YES	Work done in the context of WP3: - Apply adversarial attacks to AI models and report performance on edge devices. - Defence methods developed in the context of Use Cases