International standards portfolio to enable responsible and certifiable AI systems

Marta Janczarski
Susanne Kuch
Kimberly Lucy
Paul Cotton
Peter Deussen
AI that serves humans: demand for oversight and control

- With evolution of IT-Infrastructure in business and daily life raises the demand for secure, reliable and trustworthy IT systems and devices.

- We already deploy highly capable IT systems, software and AI applications today.

→ The demand for a risk manageable, safe, trustworthy and reliable AI is therefore a necessary evolutionary step (see also Future of Life Institute open letter).
International Standards for controlled, responsible and trustworthy AI

• 20 years of experience for secure, reliable and robust IT systems to build on

• Development of AI specific international standards help to fulfill the demands for safe, responsible and trustworthy AI
Key principles for creating standards that enable responsible and certifiable AI systems

**Flexible**
- to adapt to different organizations with different needs across the world

**Durable**
- scoped to responsible behaviors rather than highly specific technical content given how fast the industry changes

**Transversal & Scalable**
- built on the work of existing tech standards ecosystems (e.g., 27000 series for security)
What might a responsible AI standards ecosystem look like in an organization?
ISO/IEC 42001, Information technology — Artificial intelligence — Management system (AIMS)

Abstract

This document specifies the requirements and provides guidance for establishing, implementing, maintaining and continually improving an AI (artificial intelligence) management system within the context of an organization. This document is intended for use by an organization providing or using products or services that utilize AI systems. This document is intended to help the organization develop, provide or use AI systems responsibly in pursuing its objectives and meet applicable regulatory requirements, obligations related to interested parties and expectations from them.

- Standard suitable to be audited and certified by 3rd party auditors
- Contains:
  - Requirements on how to implement processes for managing AI systems in a responsible way, including
    - risk management
    - AI system impact assessment (individuals, society)
  - Extensive operational/system level controls and implementation guidance to help organization to meet these requirements
- Expected to be published end 2023

ISO / IEC AI Workshop | 12, 14 June 2023

SC 42 – Artificial Intelligence
ISO/IEC 42005, Information technology — Artificial intelligence — AI system impact assessment

Abstract

This document provides guidance for organizations performing AI system impact assessments for individuals and societies that can be affected by an AI system and its foreseeable applications. It includes considerations for how and when to perform such assessments and at what stages of the AI system lifecycle, as well as guidance for 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.

- Guidance document
- Contains:
  - Recommendation on how to implement an AI system impact assessment process
  - Involve stakeholders
  - Document results
  - How to integrate with ISO/IEC 42001 and ISO/IEC 23894:2023
- Expected to be published 2025
Abstract

This document provides guidance on how organizations that develop, produce, deploy or use products, systems and services that utilize artificial intelligence (AI) can manage risk specifically related to AI. The guidance also aims to assist organizations to integrate risk management into their AI-related activities and functions. It moreover describes processes for the effective implementation and integration of AI risk management.

- Guidance document, based on ISO 31000:2018
- Contains:
  - Principles for risk management
  - Framework for the set-up of risk management processes and integration into other organizational processes
  - Processes:
    Scoping in the context of the organization, risk assessment, risk treatment, communication and consultation, monitoring and review, recording and reporting
ISO/IEC 42006
Information technology — Artificial intelligence — Requirements for bodies providing audit and certification of artificial intelligence management systems

Abstract

This document specifies additional requirements for ISO/IEC 17021-1 in order to enable accredited and or peer assessed certification bodies to reliably audit the management system for organizations that develop or use AI systems or both according to ISO/IEC 42001 and to make an evaluation and decision for certification. The application of this document enables the certification bodies to meet the specific technical features and the particular risks in dealing with AI systems according to ISO/IEC 42001. This allows accreditation bodies and peer assessors to assess the competencies of conformity assessment bodies in an efficient and harmonized way and ensures the comparability and reproducibility of certificates confirming conformity with ISO/IEC 42001. The requirements contained in this document shall be demonstrated by any body providing AIMS certification.

- Specifies ISO CASCO Standard ISO/IEC 17021-1 for conformity assessment of management systems
- Framework for third-parties to assess correct implementation of AIMS by organizations according to ISO/IEC 42001
- Defines qualification levels and relevant knowledge for AIMS auditor teams
- Expected to be published in 2024
Other supporting standards in the foundational ecosystem

- 38507 (Gov. Implications of use of AI)
- 23894 (Risk Management)
- 5259 series on data quality
- 42001 MSS Requirements and Guidance
- 22989 on AI concepts and terminology
- 23053 Framework for AI systems using machine learning
- 42006 (Certification Body requirements)
- Certification/Conformance
- Customer Trust Accountability Transparency
Next Steps

- Implementation of Ad hoc Group (AHG) on applying AIMS in various sectors (verticals)

- Building bridges and awareness across the MSS family

- Using experiences and applications of AIMS:
  - Standards Council of Canada Pilot on Accreditation
  - Adoption by CEN-CENELEC as a European standard

- Further develop all standards needed for a responsible, trustworthy AI

- Standardize tools for testing and assessing a responsible and trustworthy AI, where necessary, that can also be used for conformity assessment
Summary

- Experts in SC 42 work on various standards to achieve objective requirements for developing and applying a responsible and trustworthy AI.

- Experts in SC 42 are focused on creating standards that intersect with existing or developing cybersecurity and IT standards.

- Stakeholders of all interested parties are present and further welcomed to participate via their National Bodies.
Thank you

Marta Janczarski (mjanczarski@microsoft.com)

Susanne Kuch (susanne.kuch@dakks.de)

Kimberly Lucy (kilucy@microsoft.com)

Paul Cotton (paul.cotton@outlook.com)

Peter Deussen (Peter.Deusse@hotmail.com)