AI Security Standardization

ISO/IEC 42001 Certification and AI Security: SGS's Expertise and Lessons Learned

Michal Cichocki | 2025-05-08 I NASK Workshop “AI Security Standards”

SGS: The World TIC Leader

Testing

2 500
Offices & labs

Inspection

No.1
World Leader

Certification

99 500
Employees

Our history

1878
SGS is founded

Mid-20th century
Diversified into inspection, testing and verification services

1981
Listed on the Swiss Stock Exchange

Today
145+ years in business

SGS: The World Leader in AI Assurance Services

Largest Network

500+ digital trust experts and 10 accredited cyberlabs

First

with 2 accreditation for AI management system certification against ISO/IEC 42001

ANAB
ANSI National Accreditation Board
ACCREDITED
ISO/IEC 17021-1
MANAGEMENT SYSTEMS CERTIFICATION BODY

SAC Accredited Certification Body

First

to deliver AI management system certification

First

to deliver Common Criteria evaluation of AI

First

to deliver Common Criteria evaluation of Post-Quantum Cryptography

SGS

Need for AI Security Standards

- Rapid AI adoption and increasing security risks
  - Generative AI as a tool for malicious actors
  - Exploitation of deepfakes for misinformation
  - Risks of rushed AI adoption without security considerations
  - Lag in AI regulation creating opportunities for self-regulation
- Need for a foundational framework through standardization
  - Building trust and confidence in AI technologies
  - Facilitating regulatory compliance
  - Customer demand for assurance/trust of AI systems
  - Improving risk management and governance
- ISO/IEC 42001: First international management system standard for AI (December 2023)

Navigating the Challenges of AI Security Standardization

- Absence of universal definitions and terminology
- Rapid evolution and technical complexity of AI
- Balancing governance with innovation
- Data privacy and security concerns
- Unique AI-specific security risks
  - Adversarial Attacks
  - Data Poisoning
  - Model Theft
  - Prompt Injection Attacks
- Integration with existing cybersecurity infrastructure
- Need for multidisciplinary expertise and collaboration
  - SGS experts participate in working groups ISO/IEC JTC SC42, SC 27 and CEN-CENELEC JTC 21

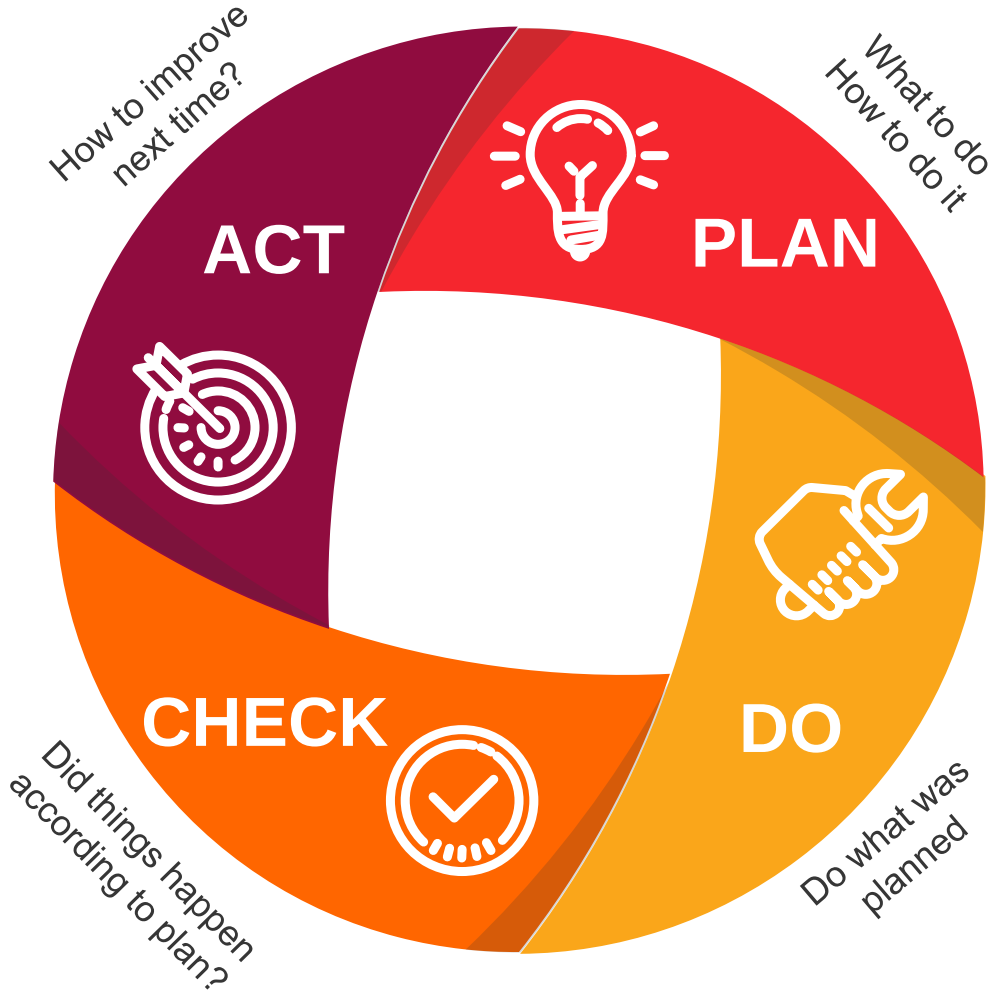
ISO/IEC 42001

It specifies requirements and gives guidance on establishing, implementing, maintaining and continually improving an AI management system.

ISO/IEC 42001 can help organizations develop or use AI systems responsibly in pursuit of their objectives while meeting regulatory requirements, as well as the obligations and expectations of interested parties.

Key requirements

- Continuous improvement
- AI risk management
- AI Controls (Annex A+B)
- AI impact assessment



ISO/IEC 42001 certified clients (selection)

AI Clearing

Poland/USA

AI-powered progress tracking with QC reporting for construction projects

Godot

Japan

AI-driven solutions for behavioral gap analysis and personalized communication

SPAN

Croatia

Development of custom AI solutions

Changi Airport

Singapore

AI-driven applications for passengers

Xayn / Noxtua

Germany

Europe’s first sovereign Legal AI (LLM)

OrionStar Robotics

China

AI System (LLM) supporting robots

Challenges and Considerations for ISO/IEC 42001 Adoption

- Complexity of understanding and implementing requirements
- Newness of the standard – limited prior experience
- Initial investment of time and resources
- Ongoing commitment to continuous improvement and audits
- Integration with existing management systems
- Using professional GRC software/platforms
- Risk assessment conducted by a multidisciplinary team of professionals

- Implementation of proper controls to mitigate the risks (not only from Annex A)
  - Using a Data Center with a very high level of security and availability – wide cooperation
  - Employees' consent to be questioned in the event of a data leak
  - Using professional suppliers monitoring infrastructure security
  - Semi-automated pre-deployment tests of AI models covering performance benchmarking, transparency, fairness, security (people have to be involved)

Perspectives on the Impact of ISO/IEC 42001 on AI Security

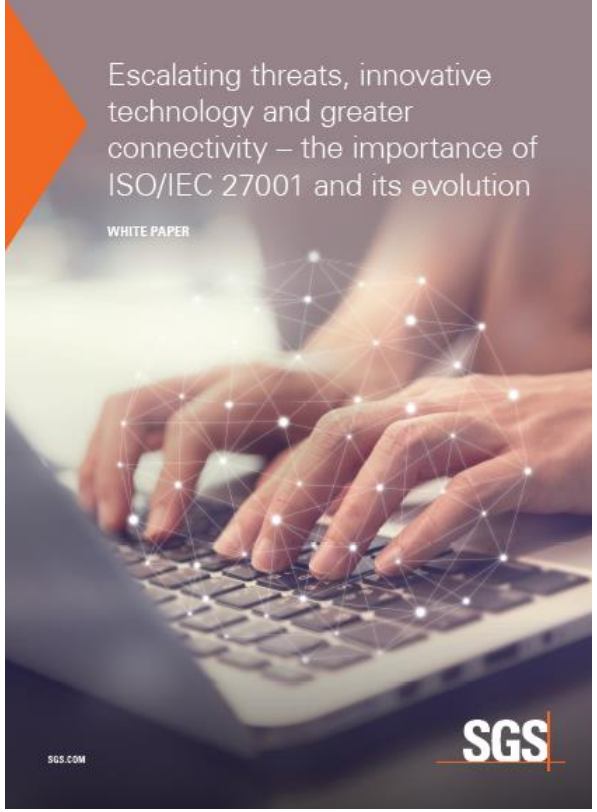
- Holistic risk management framework with a comprehensive set of supporting standards
- Alignment with other security standards and regulations
  - ISO/IEC 27001
  - EU AI Act / NIST AI RMF
- Promoting responsible AI development and deployment
- Helps meeting coming regulations and conformity assessment
- Additional ISO/IEC 42006 requirements for certification bodies
  - Requirements for auditors / Ongoing professional development
  - Pre-certification activities (scope, AI roles, audit time)

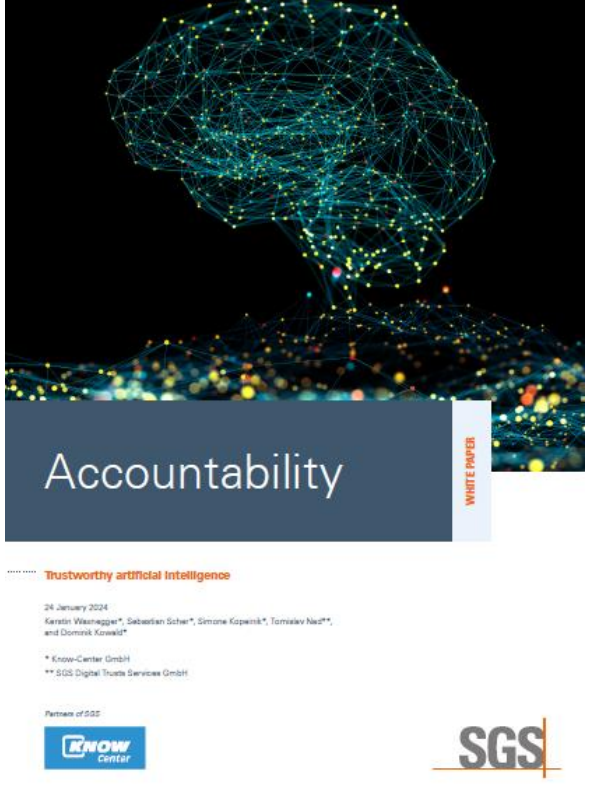
Recommendations for Organizations

- Integrate security and ethics early (in development process)
- Consider adopting ISO/IEC 42001 and other coming standards (also for EU AI Act)
- Stay informed about standards and threats
- Invest in training and expertise
- Foster collaboration between AI and cybersecurity teams
- Ensure AI Security with Robust Cybersecurity Measures

Recommendations for Standardization Bodies

- Continue developing and refining standards (including more specific standards for AI applications)
- Focus on AI-specific threats
- Emergence of sector-specific standards
- Facilitate stakeholder collaboration (Industry consortia like Cloud Security Alliance)
- Continuous adaptation is necessary due to the rapidly changing AI and cybersecurity landscape – shorter revision cycle





Further reading and recordings on SGS.com

A selection of case studies and white papers:

- AI Clearing Receives SGS's First ISO/IEC 42001 Certificate
- Escalating Threats, Innovative Technology and Greater Connectivity – the Importance of ISO/IEC 27001 and Its Evolution
- Transparency and Explainability in AI
- AI Accountability

Illuminate your AI journey with ISO/IEC 42001 Ensuring transparency, ethical management, and robust security

Do you have any questions?

michal.cichocki@sgs.com

Digital Trust Assurance

When you need to be sure