

Standards for inclusive and responsible artificial intelligence (AI)



AI is rapidly becoming central to the global economy. **Different forms of AI** – including, more recently, generative AI (GenAI) – **are already transforming the way companies, governments and communities transact and interact**, contributing to the economic dynamism of economies.

There are many examples of innovative and transformative use cases for AI across ASEAN Member States. Among them include:

- **Inclusive crisis response and recovery:** During the COVID-19 crisis, many ASEAN Member States made the use of mobile contact-tracing applications an effective means to control the impact of the pandemic. Brunei had integrated AI into their health application to assist with prediction and planning, while Thailand deployed an AI-assisted patient screening system for COVID-19 to shorten waiting times in diagnoses.
- **Natural disaster readiness and resilience:** Many AMS are vulnerable to natural disasters, particularly changing weather patterns, climate change, and consequently the natural disasters that arise from these changes such as floods and droughts. Cambodia's early warning system uses a network of smart sensors and systems that automates warning calls and messages when an impending flood is detected.
- **Promotion of gender equality, disability and social inclusion:** AI has the potential to help improve access and participation through assistive technology and rehabilitative health services. It can collect data to identify and address the needs of vulnerable and marginalised groups, particularly those that may not necessarily be visible in government databases. This includes collecting sensitive data which may be difficult to collect (e.g., data indicating violence or abuse). A hospital in Indonesia has been trialing the use of AI as assistive technology to provide supported communication options for people who are non-verbal, while a start-up in Malaysia has developed an AI surveillance system that automatically detects patient falls through AI-powered cameras placed in hospital rooms, which could be beneficial for supporting older persons.
- **Inclusive trade and SMEs:** Logistics is an important component of the trade process, with the first- and last-mile delivery process often being the costliest part of the supply chain. A start-up in Viet Nam has developed an AI-powered route optimization and transport management platform, which in turn helps to increase speed and reduce cost.

Key challenges and risks

While there are benefits to GenAI in driving innovation and efficiency, it is also paramount to ensure fair, secure and responsible use of AI to mitigate potential challenges and risks around ethics, rights and safety, particularly to vulnerable groups such as women and children.

- **Data protection and privacy:** GenAI models train on large datasets, which may involve personal or sensitive data. If there are inadequate data security measures, there is a risk of breaches, theft, or misuse of such data, leading to potential infringement of individuals' privacy.
- **Misinformation and manipulation:** Synthetic content generated by GenAI may be used by malicious actors to spread misinformation, manipulate public opinion, or launch social engineering attacks. Significant impact to political stability, economic activity, or social order may result.
- **Bias and fairness:** Training data in GenAI models may contain biased or inaccurate information, which can generate content that influence decisions that reinforce stereotypes or discrimination. This can be problematic when applied to job recruiting, risk assessments, and law enforcement.
- **Intellectual property (IP) rights:** There are concerns about the use of copyrighted materials by GenAI models without the proper authorisation from copyright holders, which would lead to infringements upon the rights of the copyright holders. The reproduction of copyrighted material for the purposes of making content accessible for people who are blind, visually impaired or otherwise print-disabled is, however, permitted under the Marrakesh Treaty.

Key standards for the use of AI

Data protection and privacy

- The ISO/IEC 27000 family of standards is widely regarded as a baseline for data protection principles (ISO/IEC 27001 on Information Security Management and ISO/IEC 27002 on codes of practice for information security controls).

Data quality, ethics and management

- ISO/IEC SD 5259 describes a data quality model for data analytics and AI based on machine learning.
- IEEE P2247.1 governs the classification of AI-driven adaptive instructional systems.
- IEEE P2660.1 covers practices for the integration of low-level automation functions and other software agents in industrial control platforms.

Subcommittee 42 (SC 42) of ISO/IEC JTC 1 oversees the development standards around AI ethics such as bias and transparency.

- IEEE 7000-2021 addresses ethical concerns during system design, among others.
- ISO/IEC TR 24028:2020 on the trustworthiness of AI, covers issues like transparency, while
- ISO/IEC TR 24027 addresses bias in AI decision-making.
- In terms of inclusion and accessibility, ISO/IEC GUIDE 71:2014 addresses accessibility in standards, ISO/IEC 30071-1:2019 covers the accessibility of user interfaces, and ISO 21801-1:2020 tackles cognitive accessibility

Awareness-raising and capacity-building among all stakeholders (individuals, businesses, governments) are essential to promote wider adoption of standards, including those that can help to mitigate risks, particularly to vulnerable groups.

ASEAN Guide on AI Governance & Ethics

ASEAN published a [Guide on AI Governance and Ethics](#), endorsed at the 4th ASEAN Digital Ministers Meeting (ADGMIN), in February 2024.

The Guide aims to **promote consumer confidence and facilitate cross-border deployment** of AI-powered services and solutions, enabling ASEAN to fully leverage the power of AI as a region. It focuses on encouraging alignment within ASEAN and fostering the interoperability of AI frameworks across jurisdictions.

The AI governance framework is underpinned by **seven guiding principles** to ensure trust in AI and the design, development and deployment of ethical AI systems, while considering the broader societal impact:

- Transparency and explainability
- Fairness and equity
- Security and safety
- Robustness and reliability
- Human-centricity
- Privacy and data governance
- Accountability and integrity

The AI governance framework has **four key components** to promote the responsible use of AI:

- Internal governance structures and measures
- Determining the level of human involvement in AI-augmented decision-making
- Operations management
- Stakeholder interaction and communication

It also details specific use cases by organisations in line with the AI governance framework.

For more information, kindly visit:

<https://asean.org/book/asean-guide-on-ai-governance-and-ethics/>

