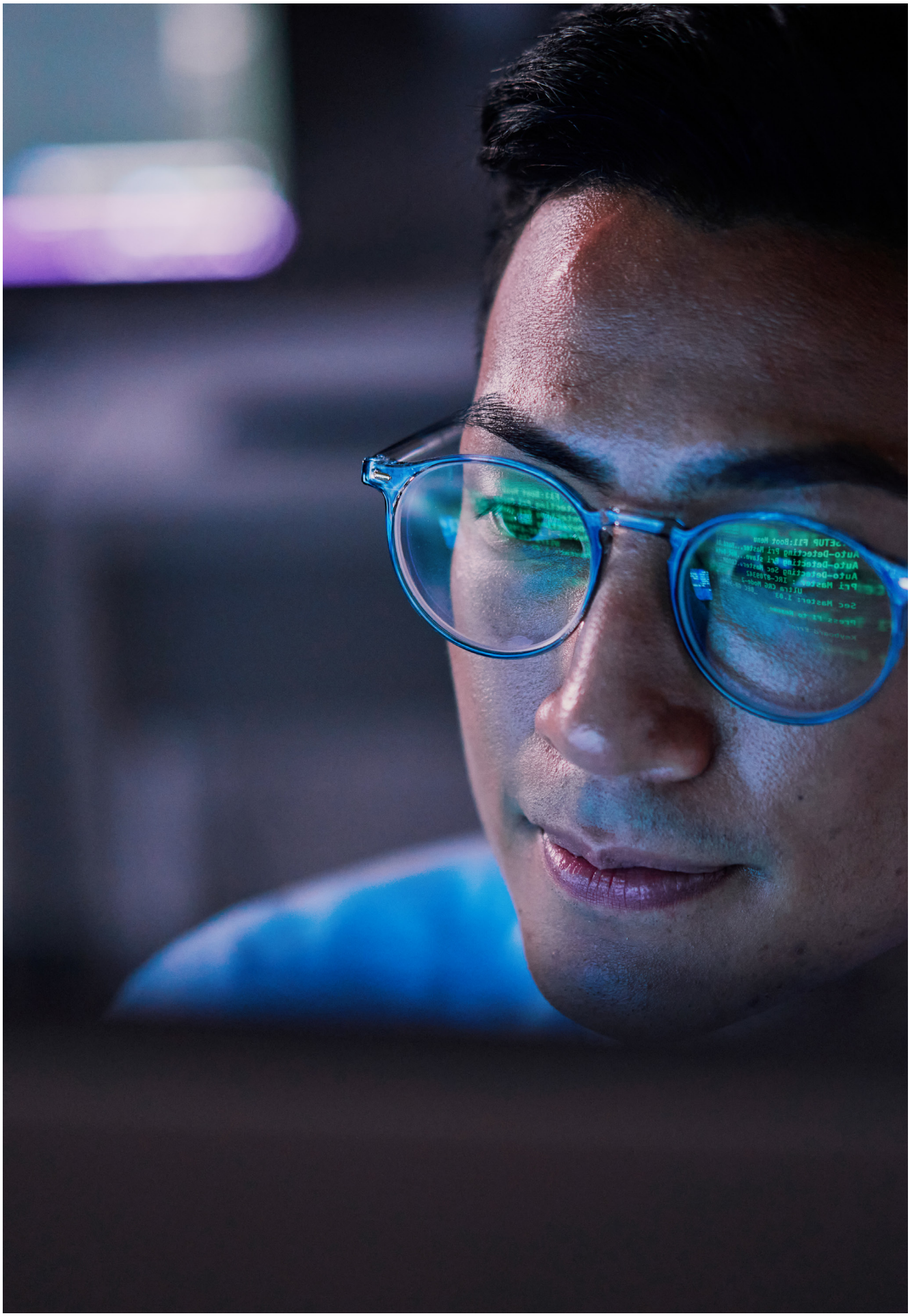
ISO policy brief: Harnessing international standards for responsible AI development and governance

iso.org



Contents

Executive summary

Artificial Intelligence is rapidly transforming global economies and society, creating unprecedented opportunities for economic growth, scientific advancement, and sustainable development. At the same time, it presents complex challenges spanning from bias, intellectual property, and privacy to effective human oversight and organizational risk management. As AI technologies and applications advance in complexity, sophistication, and generality, internationally-aligned best practices and governance frameworks that consider both technology and its interaction with people become increasingly important.

This policy brief describes the unique and valuable roles that international standards play in supporting responsible AI development and governance. International standards:

- Establish a common language and consensus-built definitions that accelerate innovation by enabling more productive collaboration among AI developers, deployers, governments and regulators, and other important stakeholders.
- Set out consensus-driven metrics, benchmarks, and technical requirements that can facilitate transparency, consumer choice, and trade, while remaining adaptable to the diverse contexts in which AI systems are deployed.
- Translate high level principles for responsible AI into concrete, actionable steps and technical requirements, supporting effective implementation of responsible AI frameworks.
- Offer detailed specifications and guidelines that can be used by regulators to improve the technical rigor and international interoperability of AI-related regulation, improving governance in a way that facilitates trade and eases compliance for AI developers.
- Underpin robust conformity assessment procedures that enable verification of technical and organizational requirements, helping to improve the reliability, quality, and trustworthiness of AI systems.

In short, international standards provide a technical foundation for advancing trustworthy AI innovation and governance.

Completed and in-progress standardization projects within the ISO and IEC system address a variety of technical and management topics relevant to the entire life cycle of AI systems. These specific AI standardization projects:

- Help ensure quality, safety, and reliability of AI products and services by providing frameworks, methods, and metrics for evaluating quality and robustness, treating unwanted bias, and mitigating AI-related cybersecurity threats (see Section 4.1);
- Improve data quality by providing principles, frameworks, and guidelines for assessing data quality, establishing data governance mechanisms, and protecting data privacy (see Section 4.2);
- Establish robust AI governance structures by providing consensus-driven risk management frameworks, overviews of ethical and societal concerns, and methods and taxonomies that improve AI transparency and interpretability (see Section 4.3);
- Advance innovation and economic growth by providing common and interoperable AI-related definitions, concepts, and terminology (see Section 4.4);
- Advance sustainability by setting out definitions, methodologies, indicators, and frameworks that can enable rigorous measurement, reporting, and management of AI systems' environmental impacts (see Section 4.5);
- Support compliance by underpinning robust conformity assessment (Section 4.6):
- Support responsible applications of AI through sector-specific metrics, guidance, and frameworks (see Section 4.7).

The global policy landscape reflects a growing recognition of the importance of international standards to AI governance. From international initiatives such as the United Nations' Global Digital Compact, to recent AI legislation such as the European Union's AI Act, to various regional and national AI strategies, policymakers increasingly recognize the need to incorporate and leverage international standards to promote good governance mechanisms, encourage interoperability and aid compliance, accelerate innovation, protect consumers, and enable robust conformity assessments.

The opportunities, risks, and impacts of AI do not emerge from technical systems in a vacuum – rather, they depend on the complex interactions between technology, people, and society. As such, considering the *socio-technical* context in which AI systems will be deployed is important to make AI standards effective in a wide variety of settings, responsive to market needs, and impactful. International standards development organizations provide a structured, neutral platform for the interdisciplinary collaboration required to achieve this balance - bringing together technical experts and experts in the broader contexts of AI systems' design and deployment, including policymakers, regulators, consumers, and civil society to integrate technical realities and societal considerations.

As AI technologies and application contexts continue to evolve, international standards can provide a robust foundation for responsible AI innovation that serves the global public interest. Strengthened collaboration between standards development organizations, national standards bodies, governments and regulators, and civil society can help ensure that AI's transformative potential benefits people around the world while minimizing its risks.

Glossary of abbreviations

Abbreviation	Definition
AI	Artificial Intelligence
ASEAN	Association of Southeast Asian Nations
AU	African Union
CASCO	ISO Committee on Conformity Assessment
GDC	Global Digital Compact
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
NSB	National Standards Body
OECD	Organisation for Economic Co-operation and Development
OHCHR	United Nations Office of the High Commissioner for Human Rights
SDO	Standards development organization
UN	United Nations
UNESCO	United Nations Educational Scientific and Cultural Organization
UNGCP	United Nations Guidelines for Consumer Protection
UNGPs	United Nations Guiding Principles on Business and Human Rights
WTO	World Trade Organization

1. Introduction

The ISO thematic policy brief series consists of concise documents designed to support ISO members (National Standards Bodies) and policymakers in their engagement to address key policy issues. Each brief offers a global perspective that brings together policy and standards insights, fostering shared understanding and encouraging coordinated action toward common goals.

The ISO thematic policy brief series explain why an issue is a policy priority, highlight the challenges involved, and provide an overview of how governments are responding at the international, regional and national levels. They also outline the role of international standards and National Standards Bodies (NSBs) in supporting policy responses, and signpost key international standards, supported by real-life examples, that can help policymakers achieve a range of policy objectives. Each document concludes with recommendations as well as suggestions on how to stay abreast of ISO's technical work.

The ISO policy briefs should be read alongside both the ISO Standards and Public Policy toolkit and the ISO policy brief: A primer on public policy – Maximizing your NSBs engagement with policymakers, which explain the pivotal role of NSBs in supporting policy across various stages of the public policy life cycle - from problem identification and policy analysis, through policy formulation and stakeholder engagement, to and implementation).

One of these thematic publications is this ISO policy brief: *Harnessing International Standards for responsible AI development and governance*. In this brief, "AI development and governance" is understood to encompass the full lifecycle of AI systems, from from conception and design, through development, deployment and operations, to governance and retirement. This policy brief is intended to be a living document, particularly given the pace of development and adoption of AI technologies, and will be periodically updated.

Acknowledgement

This publication was developed by the ISO Central Secretariat (ISO/CS), with overall leadership provided by Cindy Parokkil of the Capacity Building Unit, and with oversight from Erich Kieck, Director of the Unit.

The content was prepared by Cindy Parokkil, Matt O'Shaughnessy and Belinda Cleeland.

The publication has greatly benefited from the peer review by:

- British Standards Institution (Sahar Danesh)
- Bureau of Indian Standards (Kshitij Bathla)
- Egypt Ministry of Communications and Information Technology (Nevine Tewfik)
- ISO/IEC JTC 1/SC 42 leadership team
- OECD.AI team
- Standards Australia (Kyle McCurdy)
- UN ODET (Mehdi Snene)
- UN OHCHR (Michel Roberto Oliveira De Souza)
- World Bank Group (Sharmista Appaya Mandepanda)
- World Trade Organization (Emmanuelle Ganne, Lauro Locks, Anastasiia Koltunova and Erik Wijkstrom)

The views expressed in this brief do not necessarily reflect those of the reviewers. Any errors remain the authors' own.

We also thank the ISO/CS Marketing and Communication team for their support in finalizing the publication.

This work was made possible through funding from Korean Agency for Technology and Standards (KATS), within the framework of the ISO Action Plan for developing countries. We gratefully acknowledge KATS's ongoing and generous support for ISO's initiatives in developing countries.

Disclaimer

This publication does not reflect the opinions of ISO concerning the legal status of any country, territory, city or area or its authorities, or the delimitation of its frontiers or boundaries. The views and recommendations contained herein are solely those of the author and do not reflect the official stance of ISO or its affiliates. While every effort has been made to verify the information in this document, ISO cannot accept responsibility for any errors it may contain. This document serves strictly as an informational resource and does not represent the consensus views found in ISO standards and other ISO deliverables.

2. AI Governance is a global challenge

Artificial Intelligence (AI) technologies are increasingly being embedded in our daily lives, rapidly transforming the ways people live and work. From recommendation algorithms to healthcare diagnostics, AI is becoming increasingly prevalent and impactful. To harness its potential while minimizing risks, careful and responsible governance is essential across the entire AI life cycle.

AI technologies have the potential to drive progress on some of the world's most pressing challenges – for example, helping to combat climate change, strengthen food security, and boost economic growth. However, without careful and responsible governance across the AI life cycle – including its design, development, and deployment – AI technologies could also lead to serious individual and societal harms, such as by harming consumers, exacerbating inequalities, and threatening the enjoyment of human rights, including privacy.

At their core, AI systems are socio-technical in nature: they do not operate in isolation, but interact continuously with people, institutions and processes, cultural norms, other technologies, and broader social, economic and political contexts. Effective application and governance of AI technologies must account for this complex web of interdependencies -- which requires an understanding not only of the technical dimensions of an AI tool, but also of how it integrates and impacts the wider ecosystem in which it operates. This has been emphasized by both civil society and industry stakeholders:

"Applying a *sociotechnical* lens to AI governance means understanding how AI-powered systems might interact with one another, with people, with other processes, and within their context of deployment in unexpected ways...These approaches add a wealth of insight to teams developing AI-powered systems, helping technologists understand how users interact with products, how technologies affect social groups and economies, and how the impacts of technology can emerge over time as AI systems and people co-evolve. "

– Center for Democracy and Technology

"This [socio-technical] approach is rooted in the observation that AI systems are sociotechnical systems: both humans and machines are necessary in order to make the technology work as intended - Consequently, AI evaluation requires a framework that integrates these components and their interactions."

– Google Deepmind

Others have explicitly called for the incorporation of socio-technical considerations in standards. For example, the UN High-Level Advisory Body on AI's final report, "Governing AI for Humanity," states that "...as AI technologies increasingly impact society, there is a need to shift to a socio-technical paradigm. This shift acknowledges that AI systems do not act in a vacuum; they interact with human users and affect societal structures. Modern AI standards can integrate ethical, cultural and societal considerations alongside technical requirements."[1]

The global impacts of AI technologies add another layer of complexity. While cutting-edge AI development is often concentrated in only a few countries, its impacts are nevertheless felt worldwide, reshaping international trade, labour markets, and geopolitical dynamics. Effective AI governance, therefore, requires a nuanced understanding of the many global contexts in which AI technologies may be used, as well as the broad global impacts AI technologies can create. International cooperation is particularly crucial to build this nuanced understanding and ensure that AI technologies benefit people around the world, including in developing countries.

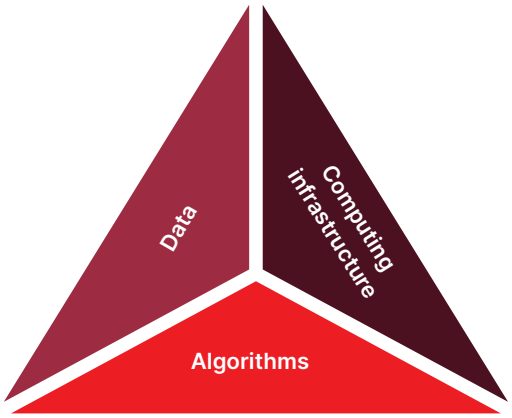
The UN Global Digital Compact (GDC), adopted in September 2024, represents a detailed commitment by UN Member States to the importance and relevance of AI standards and international collaboration in AI governance. It calls for standards development organizations (SDOs) to "collaborate to promote the development and adoption of interoperable AI standards that uphold safety, reliability, sustainability and human rights" (paragraph 58).

1 https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf

2.1 What is AI?

To understand the role of standards in AI development and governance, it is essential to establish a clear understanding of the breadth of AI technologies for which standards are relevant. Although many varying definitions of AI exist in technical documents and policy instruments, at its core, AI refers to computer systems able to perform tasks that normally require intelligence. ISO/IEC 22989:2022 defines AI as an "engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives" and contains extensive definitions of related terminology. It is often difficult to develop precise definitions of AI and related terminology, and varying definitions are in use across policy and technical instruments.

AI development relies on three essential components, often called the "AI triad":



- The quality of data helps determine the AI system's performance
- Processing hardware, such as chips and interconnection hardware used to train and run AI systems
- Rules and models that process data and transform inputs into useful outputs

Categories of AI technologies

The term AI encompasses a vast and diverse range of technologies that differ substantially in complexity, sophistication and generality. Predictive algorithms have long been used to forecast market trends, patient outcomes in healthcare, provide personalized product recommendations, detect fraud in banking, or optimize traffic management, supply chains and factory operations. Recent advances in machine learning and deep learning have made these applications more accurate, robust, and efficient. AI tools used in computer vision applications include facial recognition technologies that are applied for security and authentication and rapid developments in robotics and autonomous driving technologies. Natural language processing tools enable highly accurate language translation systems. Recent advances in generative AI have enabled more general AI systems, including wildly popular chatbots and content generation products.

Developments in AI technologies in recent years have produced increasingly complex and performant AI models – models that can operate with a greater degree of agency and autonomy and perform a broader set of tasks (i.e., are more "general"). Some researchers and commenters describe the goal of AI research as moving from task-specific AI tools to "artificial general intelligence" (AGI, or a system with human-level performance across a broad set of tasks and domains) or even "artificial superintelligence" (ASI, or a system that outperforms humans across a broad set of domains), though experts debate the feasibility of these goals. Importantly, the core challenges that are considered in international standards related to AI – from unwanted bias, privacy, and reliability to effective testing, oversight, and governance – apply to AI technologies of all levels of complexity, sophistication, and generality.

2.2 Opportunities of AI

When governed properly, AI presents transformative opportunities across sectors, driving progress toward the UN Sustainable Development Goals. For example, AI technologies can provide:

Economic growth: AI technologies drive economic transformation by enhancing efficiency, reducing costs, and fostering innovation across industries. They can accelerate scientific advancement, optimize infrastructure (e.g., energy grids, traffic management), and create new markets by enabling new products and business models.

Social benefits: AI technologies can improve health through improved diagnostics and treatment, enhanced educational access and personalized learning, and support environmental sustainability by optimizing energy use and monitoring natural conditions. In agriculture, AI tools can boost food security through precision farming, improving yields and resource efficiency. By advancing accessibility technologies, AI tools can help empower persons with disabilities. In this and many other contexts, AI technologies can lead to products that enhance and enable the exercise of human rights.

Global development: AI technologies accelerate digital transformation, particularly in developing regions, and support equitable policymaking by providing data-driven insights into social disparities. It can also enhance crisis response efforts through improved resource coordination and aid distribution.

2.3 Risks of AI

AI technologies also introduce and exacerbate a variety of risks, from technical and operational concerns related to the reliability, transparency, and security of AI systems themselves, to broader concerns about the human, societal, and ethical impacts that stem from AI applications:

Technical limitations: The complex, probabilistic nature of AI models can lead to security vulnerabilities and a lack of robustness. These limitations can increase the potential for compromised user privacy, cyberattacks, and a lack of reliability – factors that create particular risks in application domains such as the control safety-critical infrastructure.

Training data: Training data that contains flaws or encodes historical bias can lead to inaccurate results, the perpetuation of discriminatory trends, and the spread of toxic content targeting marginalized groups. Research has shown that these risks are particularly acute for groups that are underrepresented in AI design and development, such as for users in the Global South. Effectively mitigating these risks requires a nuanced understanding of not only the technical aspects of a dataset and AI model, but the broader historical and social contexts from which data is collected, and in which an AI system is deployed.

Misuse: AI tools are commonly used to enable frauds and scams, or to facilitate the generation and spread of disinformation. AI systems can be used to facilitate violations and abuses of human rights, such as through mass surveillance systems that enable repression. Some researchers and AI developers worry that future AI tools might lower barriers for nonexperts to create weapons that cause significant harm.

Effective oversight and human agency: AI-powered autonomous systems used in applications such as the provision of social services, healthcare, or in high-impact military contexts raise important ethical questions and create operational challenges related to human agency and human-computer interaction. More generally, the increasing use of AI systems, and the increasing autonomy afforded to them in some application domains, creates risks related to overreliance on their outputs. These dynamics can weaken human judgement, erode individual agency in decision-making, and make effective human oversight more challenging.

Impacts on vulnerable persons: AI technologies often pose particular risks for vulnerable persons, including older persons, minors, persons with disabilities, low-income consumers, or persons with low digital literacy. For example, AI can be used to facilitate fraud and scams, or individuals may make personal health or investment decisions based on AI recommendations that are incorrect, potentially leading to significant financial losses and bodily harms.

Broader impacts on society: The environmental impact of AI systems has become increasingly significant in recent years as the amount of computational power used to train state-of-the-art models has increased.[2] Economists debate the potential impacts of AI tools on employment, with some warning that AI use could lead to major shifts in the labour market, and significant concerns have been raised about fair labour practices throughout the value chain of AI systems, such as for data enrichment work. Generative AI tools have raised significant concerns about the economic and cultural devaluation of human effort and creativity, including concerns related to intellectual property.

Inequality and digital divides: AI could create and deepen digital divides, exacerbating inequalities both within and across societies. At the global level, economic benefits are likely to be concentrated in a few wealthy nations and corporations, leaving others behind. At the national and local levels, limited access to AI expertise and to reliable infrastructure-such as internet and electricity – limits the ability of less-resourced regions to access to AI technologies and participate in global AI development and governance, further deepening socio-economic disparities. These divides are compounded by market concentration, which risks concentrating economic and political power in the hands of few corporations and a small number of technologically advanced countries.

Many real-world risks of AI stem from an overlapping combination of these factors and categories.

Box 1: The role of standards in AI and sustainability

AI tools can help to accelerate research and development advances that promote sustainability. But the energy consumed by computational hardware to train and run large AI systems has skyrocketed as AI models become larger and more widely adopted, creating growing adverse environmental and climate impacts. These demands arise in two distinct ways: the large one-time energy requirements of training models, which often involve massive computational resources, and the smaller but cumulative energy demands of inference, which, while modest on a per-use basis, add up significantly when models are deployed at scale across millions of users and applications.

2 For more details, see ISO/IEC TR 20226:2025, *Information Technology – Artificial intelligence – Environmental sustainability aspects of AI systems.*

Standards provide a valuable tool for assessing and mitigating these environmental impacts. For example, existing and in-progress standards provide harmonized methods for calculating the environmental impacts of AI models, enabling more transparent and exhaustive reporting; define robust environmental indicators that can support environmental policy; and set out interoperable methodologies for managing environmental impacts that can increase the adoption of best practices.[3]

Importantly, none of these risks and challenges arise only from the technical failures of AI systems operating in a vacuum: rather, they arise from the way AI systems are integrated as a component of larger technological systems, the ways in which they are used and governed, and the ways in which they interact with the societal contexts in which they are deployed. A socio-technical approach that considers these broader technical and social dimensions is thus required to effectively anticipate and manage risks of AI.

Coordinated international governance and standards that consider these socio-technical dimensions can help ensure that AI's growth is sustainable, ethical, and inclusive. Coordinated international governance and standards also supports equitable participation in AI innovation and governance, enabling all countries with foundational resources for developing technical guidelines, risk management infrastructure, and regulatory frameworks.

2.4 The role of standards in socio-technical AI governance

Standards play a crucial role in the AI governance ecosystem, complementing and enabling implementation of international agreements, national regulations, and industry self-governance. International standards reflect consensus-based best practices and provide a shared language that supports regulatory alignment, accelerates innovation and facilitates global trade, and while enabling application across different national, technological and organizational contexts. Standards also enable conformity assessment, a globally recognized process to demonstrating that a product, service, or system meets particular requirements. Assessing conformity against international standards ensures that results are recognized across countries, reducing barriers to trade and increasing consumer confidence and safety.

AI technologies and their applications are advancing quickly. International standards are an important tool for responding to this challenge, as they can support AI governance across a large range of use cases and stages of technology development and adoption. At early stages of research and development, standards can bring consensus on definitions across diverse communities. When moving from basic lab research to experimental prototypes, standards can define measurement protocols, benchmarks, metrics, and performance standards. During product development, standards can help define interfaces and specifications for testbeds. As technology development matures, standards can define and diffuse best practices for deployment, risk management, and improve interoperability.

3 For more detail, see Coalition for Sustainable AI, "Standardization for AI Environmental Sustainability," 2025 (https://www.sustainableaicoalition.org/roadmap-of-standardization/)

Standards are often viewed (at least until recently) as purely technical instruments related to topics such as interoperability, safety, and performance – in reality standards are most effective when they consider the socio-technical context that technologies operate in. Moreover, the technical details specified by standards can have far-reaching, and often times significant, impacts on consumers and the public/society. The need for AI standards to consider socio-technical dimensions is reflected in calls from many stakeholders.

Standardization offers a pathway to systematically consider the socio-technical dimension that is key for effective AI governance. The need for AI standards to consider socio-technical dimensions is reflected in calls from many stakeholders. For example, the German AI Standardization Strategy's recommendation 3 states: Consider humans as part of the system in all phases of the AI life cycle" and a recent U.S/NIST plan for global engagement on AI standards states "Many of these areas of standardization must account for or directly address interactions between AI systems and people and institutions. In other words, AI systems and their impacts are inherently replace with socio-technical, hinging on complex interactions between AI systems and humans." Furthermore, the Partnership on AI – a non-profit coalition that includes leading technology organizations has urged standards development organizations (SDOs) to "ensure that their processes are informed by appropriate replace with socio-technical expertise, diverse perspectives, as well as required resources."[4]

NSBs can play a key role in ensuring that AI-related standards are technically rigorous, incorporate socio-technical considerations, and consider diverse societal and market needs. This holistic approach to standards development can help ensure that international standards advance innovation while responding to broader public concerns and build global trust in AI technologies. To facilitate this, NSBs should invest in capacity building efforts that facilitate greater participation in standards-making processes from stakeholders with both technical and socio-technical insights, including representatives from government, industry, civil society, and academia, and including developing country stakeholders who bring unique expertise about global deployment contexts and impacts of AI technologies. This could involve engaging in more active outreach, refining processes to reduce barriers to the effective participation of stakeholders (including those who have not historically had deep engagement in standards-development processes), and providing financial and logistical support. Such action is critical to ensure that standards development processes are informed by multistakeholder, socio-technical, and global expertise, and is aligned with the WTO Technical Barries to Trade Agreement's emphasis on enabling as broad as possible participation in standard-setting[5].

4 https://partnershiponai.org/wp-content/uploads/2024/09/PAI_Policy-Alignment.pdf
5 See, for example, WTO, "Trading with Intelligence," pp. 68, 80-82.

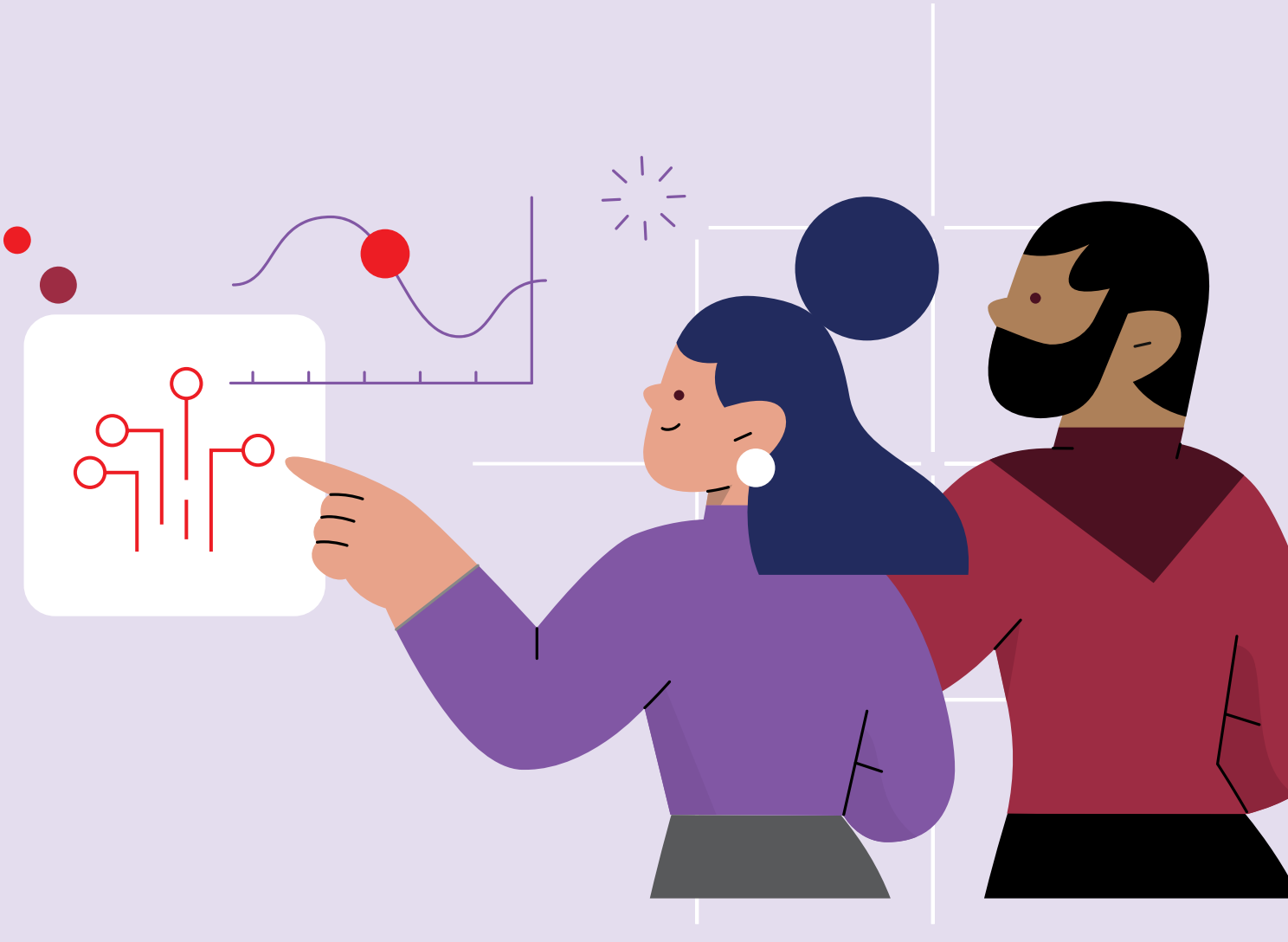
Box 2: A socio-technical approach to AI governance

A socio-technical approach recognizes that AI's impacts are shaped not just by its technological design, but also by how it is deployed in real-world contexts - how it interacts with existing processes, people, cultures, data and other technologies. For example, an AI model used for hiring may satisfy certain technical metrics for accuracy, but still exacerbate harmful bias when deployed in the real world if these accuracy metrics do not account for historical discrimination reflected in its training data. Similarly, an AI system used in healthcare may not be effective when deployed in real-world medical contexts if its designers do not consider how doctors, nurses, and other medical personnel will use it. Standards that incorporate broader social considerations can help address these challenges and ensure AI technologies benefit society, guiding AI development in a way that can more effectively implement key principles such as transparency, accountability, and fairness.

The international standardization system provides a structured, neutral space for the kind of interdisciplinary collaboration needed to foster a socio-technical approach. While technical experts bring insights into AI's technical complexities, a broad set of stakeholders, including representatives of governments, industry, consumers, and civil society, provide valuable information on the broader societal contexts in which AI systems will interact. Many existing AI-related ISO/IEC standards consider socio-technical dimensions; for example, ISO/IEC 42001:2023 (management system for AI) considers socio-technical factors such as how dynamics within organizations can affect how effectively risks are identified and addressed, and a work item on human oversight of AI systems (ISO/IEC CD 42105) considers human factors in how to provide effective oversight for autonomous systems. By integrating both technical dimensions and broader contextual dimensions, AI standards can support governance frameworks that are more effective, equitable, and innovation-promoting.

3. Policy context

Countries are using a variety of institutional mechanisms to respond to the challenges of AI governance. Many are developing AI-specific strategies or regulatory bodies, or establishing technical advisory bodies or national AI safety/security institutes. Others are developing policy frameworks that leverage existing government functions to address the impacts of AI across different domains, such as by expanding the ability of existing regulatory bodies to hire employees with AI-related expertise. International standards play a crucial role in supporting responsible AI development and governance globally.



While strategies and frameworks help articulate high-level goals and principles, concrete regulatory and oversight actions are essential for translating these ambitions into practice – this is where standards can play a critical role. Standards provide detailed, consensus-based technical and process guidance that can support implementation, monitoring, and enforcement. Effective AI governance also requires coordination across national, regional, and international levels.

Many countries have adopted national AI strategies or policy frameworks to guide the development and governance of AI technologies. These typically set out priorities related to innovation, investment, capacity-building, ethics, and inclusion. For instance, Egypt has released two editions of its National AI Strategy to strengthen regional leadership and international cooperation, and Rwanda has developed a National AI Policy in partnership with international organizations to promote responsible innovation and institutional readiness. India's National Strategy for Artificial Intelligence focuses on inclusive growth and social welfare, while Brazil's AI strategy emphasizes socially beneficial and sustainable AI, supported by investments in infrastructure, skills, public services, business innovation, and ethical governance. These frameworks provide a foundation for subsequent regulatory and institutional development, and their success will depend on moving from strategy to effective implementation.

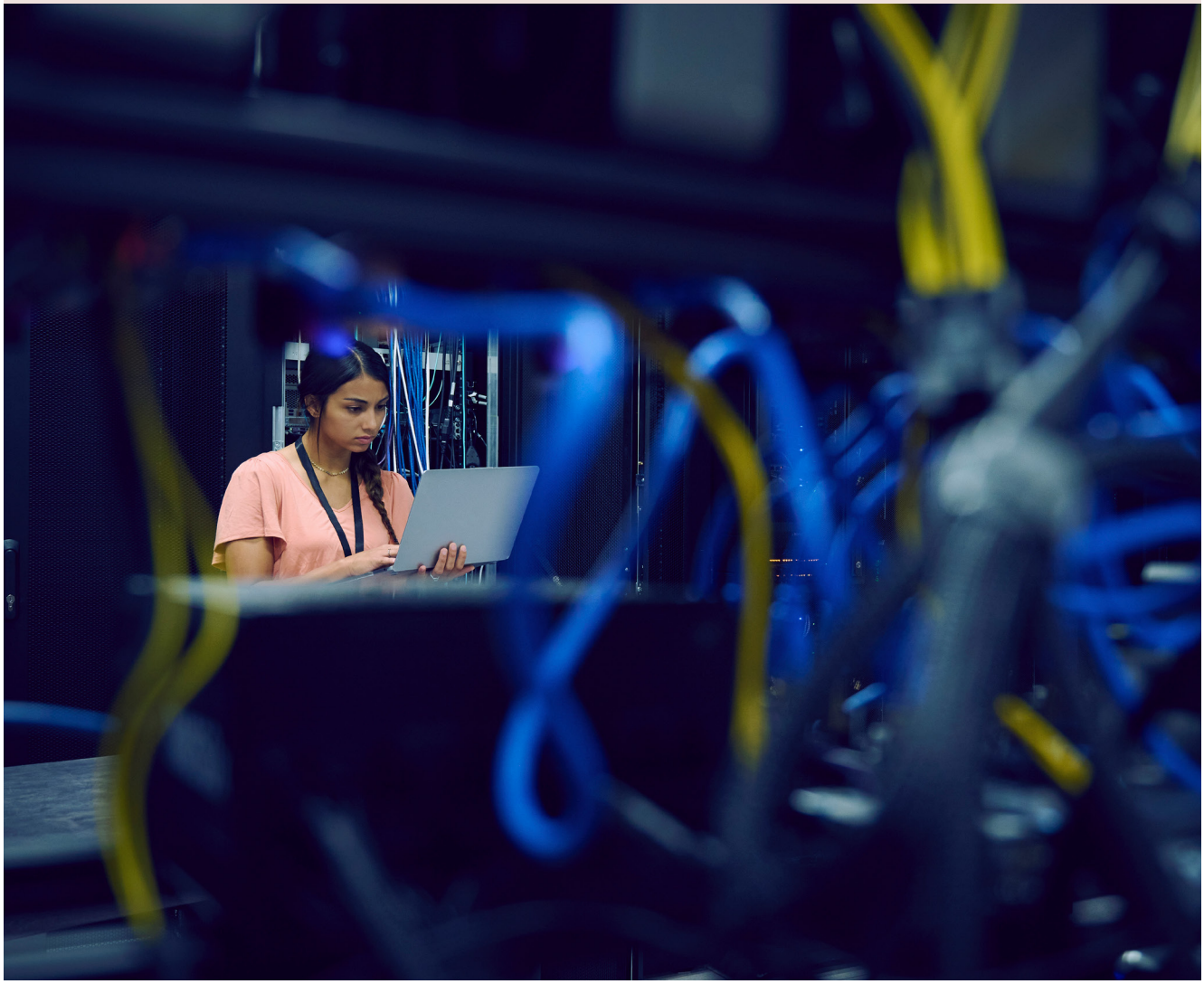
In parallel, a growing number of countries and regions are creating or empowering regulatory authorities and advisory bodies specifically tasked with AI oversight. Spain has established the *Agencia Española de Supervisión de la Inteligencia Artificial (AESIA)*, the first dedicated national AI regulatory agency in the EU. At the regional level, the *European AI Office*, established in January 2024, is responsible for overseeing implementation of the EU's AI Act. In China, the *Cyberspace Administration of China (CAC)* has assumed a central role in regulating algorithms and generative AI services, while Singapore's Advisory Council on the Ethical Use of AI and Data provides expert policy guidance to government and industry. In addition, new technical institutions have been created, such as the growing set of national AI safety and security institutes that, for example, support research on the evaluation of frontier models. International cooperation is further supported by venues such as the OECD's Working Party on AI Governance and Global Partnership on AI (GPAI), UN efforts such as UNESCO's AI ethics implementation mechanisms, and a growing set of AI-related initiatives in other international and regional bodies. The World Bank Group's 2024 report "Global Trends in AI Governance: Evolving Country Approaches" provides an overview of the emerging AI governance landscape and presents key considerations for policymakers as they shape their approaches to AI governance.

As these policies, forums, and initiatives proliferate, a lack of alignment and shared technical language between them has created fragmentation that can hinder innovation and effective governance. Such fragmentation increases compliance costs, complicates cross-border AI deployment, and raises the possibility of regulatory arbitrage, where undesirable uses or development of AI shifts to jurisdictions with weaker oversight.

International standards provide a valuable tool to mitigate this fragmentation and disseminate best practices for responsible AI development and governance. Developed through consensus-driven processes, international standards provide technical and governance benchmarks, guidance, and processes that can support interoperability while remaining adaptable to different national contexts. By integrating international standards into AI policies and regulations, governments can accelerate the shift from strategy to implementation, enhance interoperability, facilitate trade, promote best practices, and improve trust in AI systems. See section 4, for a selection of key AI-related ISO/IEC standards that can support various AI policy objectives.

Box 3: Standards, trade and good regulatory practice

International standards enable of AI governance in various ways, including by supporting good regulatory practice. This in turn can help regulations achieve their policy objectives while facilitating trade in safe and quality products. This is why international standards are promoted in various World Trade Organization (WTO) Agreements. The WTO Agreement on Technical Barriers to Trade (TBT Agreement), in particular, contains obligations strongly encouraging the use of international standards as a basis for technical regulations, standards, and conformity assessment procedures. These obligations are supported by various guidelines adopted by the TBT Committee, including the "Six Principles" for international standards - such as transparency, openness, effectiveness, and relevance. These principles provide important guidance for international standards-setters so that the standards they adopt can be a meaningful basis for governments to regulate in a way that helps avoid unnecessary barriers to trade.[6] Referencing international standards in regulation, legislation, or policy allows governments to rely on globally agreed best practices, reducing regulatory burdens and enhancing the acceptance of policies across jurisdictions. This approach also promotes regulatory cooperation, ensures regulations remain up to date independently of political cycles, and facilitates market access by aligning AI governance with international norms.



6 See WTO, "Trading with Intelligence: How AI Shapes and is Shaped by International Trade," 2024 (https://www.wto.org/english/res_e/publications_e/trading_with_intelligence_e.htm)

3.1 Selected international initiatives

In recent years, there has been a surge of international initiatives focused on AI governance, some multilateral (open to participation from all countries), others plurilateral (with participation limited to group members). Below is a selection of some influential efforts.

Organization	Policy initiative	Description	Role of standards
United Nations (UN) General Assembly	Global Digital Compact (GDC) (09/2024)	Objective 5 of the GDC pertains to "enhance[ing] the international governance of AI for the benefit of humanity." It sets out general principles for AI development and governance, and directs the creation of a multidisciplinary Independent International Scientific Panel on AI (somewhat akin to the IPCC for climate) and a Global Dialogue on AI Governance.	Paragraph 58 calls on standards development organizations "to collaborate to promote the development and adoption of interoperable artificial intelligence standards that uphold safety, reliability, sustainability and human rights"
UN Educational Scientific and Cultural Organization (UNESCO)	Recommendation on the Ethics of Artificial Intelligence (11/2021)	The Recommendation is presented as a standard-setting instrument - it defines values, principles and policies to guide countries in building legal frameworks to ensure that AI is deployed for the common good. It mandated UNESCO to produce an Ethical Impact Assessment and a Readiness Methodology to aid countries with implementation.	Standards feature prominently in the Recommendation, as tools to ensure ethical AI deployment. Member States are called upon to participate in the development of international standards for AI safety and transparency (paragraph 64) and encourage all AI actors to follow existing international standards (paragraph 72). The importance of data protection standards is emphasized.

Organisation for Economic Cooperation and Development (OECD)	Recommendation of the Council on Artificial Intelligence, often referred to as the 'OECD AI Principles' (Adopted in 2019 and updated in 2024)	The OECD Recommendation on AI is the first intergovernmental standard on AI, with 47 adherents as of August 2025. It contains 5 principles and 5 policy recommendations for responsible, human-centric AI governance. The OECD's Working Party on AI Governance (WP AIGO) supports national AI policy implementation.	Policy recommendation 2.5 states that "Governments should promote the development of multi-stakeholder, consensus-driven global technical standards for interoperable and trustworthy AI."
UN Human Rights Council (HRC) and Office of the High Commissioner for Human Rights (OHCHR)	Resolution on New and emerging digital technologies and human rights (2023) Report on Human rights and technical standard-setting processes for new and emerging digital technologies (2023)	The HRC Resolution highlights the need for States and businesses to address human rights issues related to AI and other digital technologies.	OHCHR's report on standards and human rights, requested by the Member States of the UN Human Rights Council in a 2021 resolution, emphasizes the impacts that technical standards can have on the promotion and protection of human rights, including in the context of AI, and describes recommendations for stakeholders to deepen consideration of human rights in standards-setting processes.
Council of Europe	The Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (Opened for signature on 05/09/2024)	This first legally-binding multilateral treaty on AI's objective is to ensure that "activities within the lifecycle of AI systems are fully consistent with human rights, democracy and the rule of law." It is open to non-member states of the Council of Europe, and its signatories include Canada, the EU, Japan, and the United States.	The Framework Convention sets out high-level obligations that would often be implemented through more detailed technical standards. Technical standards are explicitly mentioned in the Framework Convention's Explanatory Report as a tool that can enable implementation.

Many other international initiatives are relevant to AI governance, even if they do not focus exclusively on AI. For example, the United Nations Guidelines for Consumer Protection (UNGCP), developed by the United Nations Conference on Trade and Development sets out principles for effective consumer protection legislation, enforcement, and redress systems. These principles are supported within the ISO system by the ISO Consumer Policy Committee (COPOLCO), defining the main characteristics of effective consumer protection legislation, enforcement institutions, and redress systems. The United Nations Guiding Principles on Business and Human Rights (UNGPs) provide a framework for States and companies to prevent, address, and remedy human rights abuses in business operations, including those that involve AI.

Box 4: Human rights – A common thread in AI governance

Human rights are universal rights that belong to every person simply by virtue of being human. They include civil and political rights (such as the right to liberty and security of person, freedoms of expression and association, and the right to a fair trial) as well as social, economic, and cultural rights (such as the rights to education, work, and health). International human rights were set out in the Universal Declaration of Human Rights as well as numerous treaties, resolutions, and laws developed by the United Nations, regional human rights institutions, and governments. Widely-adopted frameworks also set out expectations for conducting human rights due diligence that are applicable to AI products and services such as the UN Guiding Principles on Business and Human Rights (UNGPs) and the OECD Guidelines for Multinational Enterprises on Responsible Business Conduct.[7]

AI technologies impact human rights in a myriad of ways. They can support assistive technologies for persons with disabilities, help to improve education, and provide new means to access information. They can also be used to facilitate censorship, undermine privacy rights, and entrench discrimination.[8] Across international, regional, and national governance efforts for AI, the protection and promotion of human rights has emerged as a key shared priority. Effectively managing the human rights impacts of AI systems requires understanding the technical and social contexts in which AI systems are designed and deployed – such as where AI systems might be used and what risks and benefits this enables. As such, socio-technical approaches to AI development and governance can help to ensure that AI activities have positive impacts on human rights.

Standards play an important role in implementing AI governance principles and frameworks, and a growing international conversation is devoted to ensuring that standards for digital technologies facilitate the enjoyment, rather than suppression, of human rights. Incorporating and deepening expertise in standards-setting processes on the potential downstream human rights impacts of standards, such as through steps to create more diverse and inclusive participation.[9]

3.2 Selected regional frameworks

The development of regional AI governance frameworks promotes cross-border harmonization, prevent regulatory arbitrage, and build consensus on key policy measures on a smaller scale. They can allow for more detailed consensus due to stronger alignment, and enable regional blocks also gain collective bargaining power - for example, through coordinated participation in standards development.

7 For details, see UN OHCHR, "Practical application of the Guiding Principles on Business and Human Rights to the activities of technology companies, including activities relating to AI," 2025 (https://www.ohchr.org/en/documents/thematic-reports/ahrc5932-practical-application-guiding-principles-business-and-human); OECD, "Recoding our understanding of responsible business conduct in science, tech and innovation," 2023 (https://www.oecd.org/en/blogs/2023/10/recoding-our-understanding-of-responsible-business-conduct-in-sc.html).

8 See, for example, UN OHCHR, "Taxonomy of Generative AI Human Rights Harms: A B-Tech Gen AI Project supplement," 2023 (https://www.ohchr.org/en/documents/tools-and-resources/taxonomy-generative-ai-human-rights-harms-b-tech-gen-ai-project)

9 For more detailed recommendations to deepen consideration of respect for human rights in standards-setting processes, see UN OHCHR, "Human rights and technical standard-setting processes for new and emerging digital technologies," 2023 (https://digitallibrary.un.org/record/4031373).

Regional actor	Policy initiative	Description	Role of standards
The European Union (EU)	The EU Artificial Intelligence Act (2024) Digital Services Act (2022) Corporate Sustainability Due Diligence Directive (2024)	The EU AI Act applies a regulatory approach to AI governance, with legal requirements focusing on the protection of health, safety, and fundamental rights. It takes a risk-based approach to AI governance by classifying AI systems into four risk levels: unacceptable risk (prohibited), high-risk (strongly regulated), limited risk (lighter obligations) and minimal risk (unregulated). It also specifies obligations for providers of general purpose AI models. Other EU legislation, such as the Digital Services Act and the Corporate Sustainability Due Diligence Directive, set out other requirements that can apply to AI-related products and services.	The AI Act, like other EU legislation, relies heavily on harmonized European standards to provide technical specifications for demonstrating compliance with its requirements. It instructs the European Commission to issue standardisation requests to the European standardisation bodies (CEN, CENELEC) to develop standards to cover the requirements of the Act (Article 40).
The Association of Southeast Asian Nations (ASEAN)	The ASEAN Guide on Governance and Ethics	This voluntary guide is aimed at organisations developing or deploying AI systems. It establishes high-level principles for trustworthy AI, taking a risk-based approach and emphasizing economic development alongside ethical considerations.	The Guide positions standards as a key governance measure and encourages organizations to refer to, and align their practices with, international standards, particularly in areas such as information security management systems, data governance, cybersecurity, and software development.
The African Union (AU)	AU AI Continental Strategy (07/2024)	The Strategy aims to equip the AU member states with the capacities necessary to optimise AI's benefits. It proposes fifteen policy recommendations across five focus areas (harnessing AI's benefits, building AI capabilities, minimizing risks, stimulating investment, and fostering cooperation).	The Strategy's first objective is to "Implement robust AI governance, regulations, standards, codes of conduct and best practices to manage AI risks and promote its growth." It positions technical standardization as one of the fifteen core action areas needed to govern AI effectively in Africa.

3.3 National approaches

National approaches to AI governance vary widely, leveraging tools as varied as:

- Industry self-regulation through voluntary commitments and guidelines
- Support for AI development and adoption, through research and development investment
- Development of regulatory sandboxes
- Creation of mechanisms for across-government coordination
- Building government capacity to approach AI, such as through new departments, teams, and institutes
- Leveraging procurement policy to shape rules and norms for AI
- Developing training and digital literacy programs
- Creating national AI strategies
- Applying or updating sectoral regulations as they apply to AI systems
- Developing AI-specific regulation

Across all of these approaches, standards serve as a tool to support both innovation and ethics by advancing regulatory alignment, facilitating trade, and encouraging the adoption of common definitions and best practices.

Country	Policy initiative	Description	Role of standards
United States (US)	National AI Initiative Act (2020) and related frameworks (e.g. NIST AI Risk Management Framework, 2023) Accelerating Federal Use of AI through Innovation, Governance, and Public Trust (2025) America's AI Action Plan (2025)	The National AI Initiative Act is exemplary of U.S. AI policy initiatives that established coordination mechanisms for AI-related efforts across the federal government, such as on research, governance, and education. The U.S. approach relies on enforcing existing laws as they apply to AI and promoting initiatives intended to support AI research and innovation by industry.	Technical standards play a central role in the U.S. approach. The United States encourages AI developers to adopt consensus standards. The National Institute of Standards and Technology's AI Risk Management Framework (a voluntary standard for AI risk management) is widely adopted.

China	New Generation of Artificial Intelligence Development Plan (2017) plus specific laws, including Generative AI Regulations (2023) Algorithm Regulation Rules (2022) Deep Synthesis Regulation (2023)	The Plan outlines China's goals regarding AI, including its aspiration to emerge as the driving force in defining ethical norms and standards. China's AI governance approach is state-driven and includes significant new AI-focused regulations. It does not have a single AI law, but multiple binding regulations targeting various aspects of AI.	Standards are integral to China's governance approach. China often uses standards as a form of soft regulation, in which standards are developed as a testing ground for policies, first introduced as voluntary, but later often becoming part of mandatory requirements.
United Kingdom (UK)	A Pro-Innovation Approach to AI Regulation, White Paper (09/2023) National AI Strategy (2021) AI Opportunities Action Plan (2025)	The White Paper builds on the UK's National AI Strategy and centres investment, innovation, and agility and coordination among regulators in an effort to address AI harms without stifling innovation.	Standards are a key pillar of the UK's AI governance strategy. The government launched an AI Standards Hub in 2022 to boost UK's influence in international AI standard-setting and to help organizations adopt and contribute to AI-related standards.
Japan	Social Principles of Human-Centric AI (2019), Governance Guidelines for Implementation of AI Principles (revised 2022) AI Promotion Act (2025)	Japan's recent AI legislation focuses fostering AI development and adoption, using a primarily soft law approach to encourage AI developers and deployers to follow best practices to manage risks. Japan's approach is highly collaborative, involving significant participation from industry, academia and civil society.	AI governance through international standards is an explicit part of Japan's national strategy. It aligns its AI policies with international standards and plays a leading role in developing them. Japan actively promotes ISO/IEC AI standards as a central tool for AI governance.

Nigeria	National Artificial Intelligence Strategy 2024	Nigeria's AI strategy aims to make the country a global leader in AI innovation through responsible, ethical, and inclusive development. Its approach relies on existing sectoral laws (e.g., the Data Protection Act, Cybercrimes Act, Copyright Act, Competition and Consumer Protection Act), but the strategy includes a roadmap towards pro-innovation governance within five years, including the intention to create a new regulatory body for AI.	The strategy consistently treats standards as critical instruments for building a trustworthy, interoperable, and ethical AI ecosystem. It emphasizes both technical and ethical standards across its pillars, objectives, and outcomes. In particular, it stresses the crucial role of data governance and data quality standards.
Kenya	Kenya National Artificial Intelligence Strategy 2025-2030	Kenya's AI Strategy showcases its intent to build dedicated AI governance structures and legislation over time, moving from reliance on existing laws to more specialized frameworks. It envisions creating a National AI Council and has proposed a Draft AI Code of Practice, developed by the Kenyan Bureau of Standards.	Standards are positioned as bridges between innovation and regulation in Kenya's AI Strategy. The Draft AI Code of Practice proposes the development of several standards to support the Strategy.
South Korea	Basic Act on the Development of AI and the Creation of a Foundation for Trust	South Korea's AI legislation aims to establish a foundation for trustworthy AI. The legislation takes a risk-based approach, creating transparency and safety obligations for "high-impact" and generative AI systems, while also establishing government bodies to support AI development and policy.	The legislation requires the government to benchmark its regulations against international standards, undertake standardization initiatives, and support the standardization of AI technology.

4. Standards in support of AI policy

International standards can play an important role in supporting national public policy development and implementation – see ISO policy brief: A primer on public policy – International Standards and NSBs for more information.

Preparing and adopting public policies is the responsibility of policymakers and regulators. NSBs can play a vital role in supporting public policy by providing expertise, guidance and resources related to technical standards and international good practices. Their involvement can help ensure that policies are well-informed, effective and aligned with International Standards.

International standards developed by ISO, IEC and ITU, jointly known as the World Standards Cooperation (WSC), are global solutions that respond to market needs and are informed by the collective knowledge of global experts through a multi-stakeholder consensus-based process. These standards encompass a broad spectrum of perspectives and expertise, including those related to social, environmental and economic concerns. They serve as guidelines and best practices that can ensure consistency, build trust and facilitate international cooperation among countries and stakeholders. International Standards developed by the WSC, are in line with the WTO TBT Committee's "Six Principles" (see Box 3). Through the use of standards, governments can advance AI safety, reliability, alignment with consensus-based high-level ethical principles, and regulatory compliance of AI systems.

However, given the time required to develop consensus-based international standards and the rapid pace of AI development and adoption, international standards have the greatest impact by establishing foundational structures that enable trust and innovation. By ensuring good governance processes and setting clear expectations for transparency, accountability, and risk management, they provide a stable foundation for AI innovation and responsible deployment at scale.

This section describes a selection of key AI-related ISO/IEC standards that can support various AI policy objectives. The list is illustrative rather than exhaustive, as ISO and IEC continue to develop additional standards relevant to AI governance and practice. While non-ISO standards are beyond the scope of this brief, external resources such as the AI Standards Hub and nascent AI Standards Exchange (see Box 7) enumerate additional relevant standards, including those developed by IEC, ISO, IEEE, and ITU.

Box 5: The ISO and IEC system

The value of the ISO and IEC system is not solely in the standards development process and governance framework, but also in its extensive global network/membership comprising 170 national members, each with one voice representing their respective country. All members are welcome to join ISO/IEC technical committees (TCs) and contribute to the International Standards development process. Standards emerging from this process carry global legitimacy and buy-in, as they represent widespread acceptance, are based on global needs, and are the result of systematic collaboration and consensus. As a result, ISO standards have worldwide credibility and, once a consensus-based standard is agreed by the members, it can be adopted nationally by all ISO members, either as-is or with adjustments to meet the national requirements and/or context. ISO standards offer a cost-effective means of disseminating knowledge across borders and are widely used by the public and private sectors.

Box 6: The role of conformity assessment procedures

Trust is a prerequisite for AI adoption, and fostering trustworthiness in AI systems is a fundamental goal of AI development and governance. Trust is built when stakeholders have confidence that AI systems operate transparently, fairly and with accountability, and demonstrate safety, reliability, and compliance with ethical and regulatory expectations. Conformity assessment procedures, including testing, certification, and auditing, facilitate trust by allowing AI developers and regulators to demonstrate that AI products and services meet specified requirements, such as those related to transparency, security, reliability, and accuracy.

Conformity assessment relies on international standards to provide common technical language and methodology, describing clear and structured requirements that allow developers to build and manage AI systems in ways that facilitate successful certification against technical specifications in a standard or regulation. Internationally recognized conformity assessment facilitates trade and enables market access.

Within the ISO system, ISO's Committee on Conformity Assessment (CASCO) supports activities related to conformity assessment procedures. Through a Joint Working Group, ISO/IEC JTC 1/SC 42 and CASCO develop and promote conformity assessment schemes for AI systems.

Box 7: External resources for AI standards

Several external resources help stakeholders learn more about the role of standards in AI development and governance, as well as identify particular AI standards to address specific needs. The **AI Standards Hub** is a UK-led initiative that provides a searchable database of standards, training resources, and facilitates collaboration to help diverse stakeholders participate in AI standards development. The **AI Standards Exchange Database**, provides a central repository of AI standards developed by several global standards bodies.

4.1 Ensuring quality, safety, and reliability

Ensuring the quality, safety, and reliability of AI systems requires robust methods for risk identification, system validation, and cybersecurity to prevent failures and misuse. International standards offer structured approaches to achieving these goals:

- **ISO/IEC 27001:2022**, *Information security management systems – Requirements*, provides a structured, risk-based security framework that protects AI systems and their data throughout their life cycle, mitigating AI-related cybersecurity threats.
- **ISO/IEC TS 8200:2024**, *AI – Controllability of automated artificial intelligence systems*, specifies a framework with principles, characteristics and approaches to improve AI systems' controllability, maintaining their reliability and alignment with human values.
- **The ISO/IEC 24029 series**, *AI – Assessment of the robustness of neural networks*, provides methodologies to assess the ability of AI systems to maintain performance under various conditions, including changes in input data, environmental factors, and potential adversarial attacks.
- **ISO/IEC TR 24027:2021**, *Bias in AI systems and AI aided decision making*, describes measurement techniques and methods for assessing bias in AI systems.

- **ISO/IEC TS 12791:2024**, *AI – Treatment of unwanted bias in classification and regression machine learning tasks*, provides mitigation techniques that can be applied throughout the AI system life cycle in order to treat unwanted bias.
- **ISO/IEC TR 5469:2024**, *AI – Functional safety and AI systems*, describes properties, risk factors, methods, and processes for AI systems in or related to safety related functions.
- **ISO/IEC TS 25058:2024**, *Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – Guidance for quality evaluation of AI systems*, provides guidance for the evaluation of AI systems using an AI system quality model.
- **ISO/IEC 25059:2023**, *Software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – Quality model for AI systems*, outlines a quality model for AI systems that is an application-specific extension to the standards on SQuaRE.

4.2 Improve data quality and establish data governance

The performance and fairness of AI systems depends on high-quality and well-governed data. Standards help ensure transparency, privacy, and integrity in data collection and use. By integrating these standards into policy frameworks, regulators can ensure AI systems operate on reliable, high-integrity datasets, reducing unwanted bias and enhancing accountability.

- **The ISO/IEC 5259 series**, *Data quality for analytics and machine learning (ML)*, provides frameworks and guidelines to ensure high-quality data throughout the AI lifecycle.
- **The ISO/IEC 38505 series**, *Governance of data, establishes principles for data governance and accountability*, including those relevant to AI systems.
- **ISO/IEC 29100:2024**, *Security techniques – Privacy framework, defines a privacy framework for personal privacy protection*, essential for complying with data privacy laws when AI systems process personal information.
- **ISO/IEC 27018:2025**, *Code of practice for protection of personally identifiable information (PII) in public clouds acting as PII processors*, provides controls for protecting personal data in cloud environments, relevant for AI systems that often rely on cloud infrastructure for data storage and processing.
- **ISO/IEC TS 4213:2022**, *Information technology – AI – Assessment of machine learning classification performance*, specifies methodologies for measuring classification performance of machine learning models, systems, and algorithms.

4.3 Promoting good governance and ethical alignment

International standards play a critical role by translating high-level governance principles such as fairness, transparency, and accountability into concrete governance and management practices that organizations can adopt, and that policymakers can reference when crafting regulations. Some standards focus on guiding organizational governance structures and responsibilities (e.g., ISO/IEC 38507 and ISO/IEC 42001), while others address ethical and societal concerns more directly, including bias, trustworthiness, and fairness (e.g., ISO/IEC TR 24368, TR 24027, and TR 24028). Together, they provide a foundation for responsible AI deployment that is both well-governed and ethically sound.

- **ISO/IEC 23894:2023**, *AI – Guidance on risk management*, provides an AI-specific risk management framework for identifying and mitigating AI-related risks, aligning with global governance efforts to mitigate unwanted algorithmic bias and enhance AI transparency.
- **ISO/IEC 38507:2022**, *Governance implications of the use of artificial intelligence by organizations*, provides governance frameworks specifically for organizational AI use, helping leadership understand their responsibilities regarding AI.
- **ISO/IEC 42001:2023**, *AI – Management system, helps organizations establish*, implement, and continually improve their AI management systems, ensuring ethical oversight and risk governance.

- **ISO/IEC 42005:2025**, *AI system impact assessment*, provides organizations a framework to understand how AI systems may affect individuals, groups or society at large across the AI system lifecycle.
- **ISO/IEC TR 24368:2022**, *AI – Overview of ethical and societal concerns*, provides a high-level overview of AI ethical and societal concerns, supporting policy alignment.
- **ISO/IEC TR 24028:2020**, *Overview of trustworthiness in artificial intelligence*, defines the characteristics that make AI systems trustworthy and surveys approaches to improving trustworthiness.
- **ISO/IEC 12792:2025**, *Information technology – AI – Transparency taxonomy of AI systems*, defines a taxonomy to assist stakeholders with identifying and addressing the needs for transparency of AI systems.
- **ISO/IEC TS 6254:2025**, *Information technology – AI – Objectives and approaches for explainability and interpretability of machine learning (ML) models and AI systems*, describes approaches and methods that can be used to achieve explainability objectives of stakeholders.

4.4 Advancing innovation and economic growth

While governance is essential, overly restrictive AI regulations could stifle innovation. Standards help balance oversight with the flexibility, facilitating responsible innovation by ensuring interoperability, streamlining compliance, and reducing trade barriers.

- **ISO/IEC 22989:2022**, *AI concepts and terminology*, creates a common language for AI development, reducing barriers to entry by standardizing terminology across industries
- **ISO/IEC 23053:2022**, *Framework for AI systems using Machine Learning*, establishes a common set of concepts for machine learning, supporting scalable and interoperable AI development.
- **ISO/IEC 5338:2023**, *AI system life cycle processes*, ensures that AI systems are developed and managed in a structured, reliable, and scalable way, enabling organizations to innovate while maintaining control over quality and performance.
- **ISO/IEC 19941:2017**, *Cloud computing – Interoperability and portability*, establishes common terminology and concepts to achieve interoperability and portability in cloud computing, which is critical for AI data processing and model training.

4.5 Advancing sustainability

Standards provide definitions, methodologies, indicators, and frameworks that can help organizations and policymakers rigorously measure, report, and manage the impact of AI systems on the environment.

- **ISO/IEC TR 20226:2025**, *Environmental sustainability aspects of AI systems*, provides an overview of environmental sustainability aspects related to the life cycle of AI systems.
- **ISO/IEC 30134 series**, *Information technology – Data centres – Key performance indicators*, provides definitions, metrics, and key performance indicators for data center resource management and efficiency, helping operators to measure, report, and optimize the environmental impacts of their facilities.
- **ISO 14001:2015**, *Environmental management systems – Requirements with guidance for use*, provides a framework for organizations to design and implement an environmental management system.

4.6 Supporting policy, regulation and compliance

As seen in section 3, policymakers are increasingly referencing international standards in their AI governance frameworks. A major driver of this approach is that standards form part of a broader quality infrastructure system – which includes **conformity assessment** – that provides reliable and trustworthy ways to verify compliance with policies and regulations. Conformity assessment is relevant for both public and private sector actors. For governments, it offers assurance that AI systems meet regulatory and technical requirements, strengthening accountability and trust in public sector use. For industry, it provides a means to demonstrate compliance with recognized good practices, market requirements, and regulations, while reducing liability risks, and signal quality and trustworthiness to customers, investors, and regulators etc – thereby creating competitive advantages in domestic and global markets.

Conformity assessment includes mechanisms such as testing, certification, and accreditation. While these processes can take different forms, their credibility, interoperability and global acceptance recognition are greatly strengthened when underpinned by international standards.

ISO provides a comprehensive set of standards (the CASCO toolbox[10]) that establish general principles and requirements for organizations providing verification or certification services, ensuring their quality and credibility. When applied in conjunction with AI-specific technical standards, these frameworks enable robust and effective evaluation of AI systems.

For example, ISO/IEC 17021-1:2015, *Conformity assessment – Requirements for bodies providing audit and certification of management systems*, ensures that organizations providing audit and certification services for management systems such as ISO/IEC 42001:2023 (the AI management system) operate in a competent, consistent and impartial manner. ISO/IEC 42006 provides requirements for certification bodies that wish to audit AI management systems, creating an extra layer of trust in the use of AI systems. And ISO/IEC 42007 will provide a high-level framework and guidance for the development and implementation of conformity assessment schemes, including certification schemes, for AI systems.

4.7 Supporting responsible applications of AI

In addition to the "horizontal" AI standards enumerated above, which provide metrics, guidelines, and frameworks applicable to AI systems across a wide variety of applications, the ISO system also develops AI standards targeted toward uses in specific sectors where AI tools are applied. Some relevant examples include:

- **Finance** (e.g., ISO/TC 68 has developed standards for financial data and risk management relevant to AI applications, and ISO/TC 307 considers applications of AI related to blockchain and distributed ledger technology)
- **Agriculture** (e.g., ISO/TC 347 considers data-driven agrifood systems)
- **Health** (e.g., ISO/TC 210 has developed standards for quality management of medical devices that use AI, and ISO/TC 215 considers health informatics)
- **Robotics and automation** (e.g., ISO/TC 22 considers the integration of AI in vehicle safety and operation, and ISO/TC 204 considers AI in intelligent transportation systems)

In addition, ISO/IEC 5339:2024 offers guidance on applying AI in various contexts, emphasizing coherent approaches to stakeholder engagement and the AI application life cycle.

10 https://casco.iso.org/toolbox.html

5. Key messages

These following key messages highlight the opportunities and risks posed by AI, the importance of inclusive and globally coordinated approaches, and the critical function of standards in shaping trustworthy, effective, and socially responsive AI systems.

1

AI encompasses a diverse range of technologies that differ in complexity, sophistication and generality. Even as AI technologies have advanced rapidly in recent year, core challenges considered in international standards related to AI – such as bias, reliability and effective organizational oversight structures – apply to AI technologies of all levels of complexity, sophistication, and generality.

2

AI presents significant opportunities (such as economic growth, societal benefits, and global development) alongside serious risks (such as bias, privacy violations, job displacement, bodily and financial harms, and the facilitation of mis- and disinformation).

3

AI governance is a global challenge requiring international cooperation due to its worldwide impacts on trade, labour markets, and geopolitical dynamics, despite AI development being concentrated in few countries.

4

International standards complement other governance tools (regulation, agreements, industry self-governance) while supporting regulatory harmonization, facilitating global trade, and distilling best practices.

5

International standards play critical roles in ensuring the trustworthiness of AI systems at all maturity levels of AI development and across a wide variety of AI application. Standards can establish data governance, enable effective risk management, and help advance sustainable development.

6

A socio-technical approach to AI governance that considers AI in the context of the broader technical and social systems in which they operate is essential: it provides a common language for considering both technical aspects and the interactions between AI systems and people. These dimensions are important because they shape the effectiveness of AI systems and AI governance. Standards can help guide AI development and governance in ways that lead to positive societal outcomes.

7

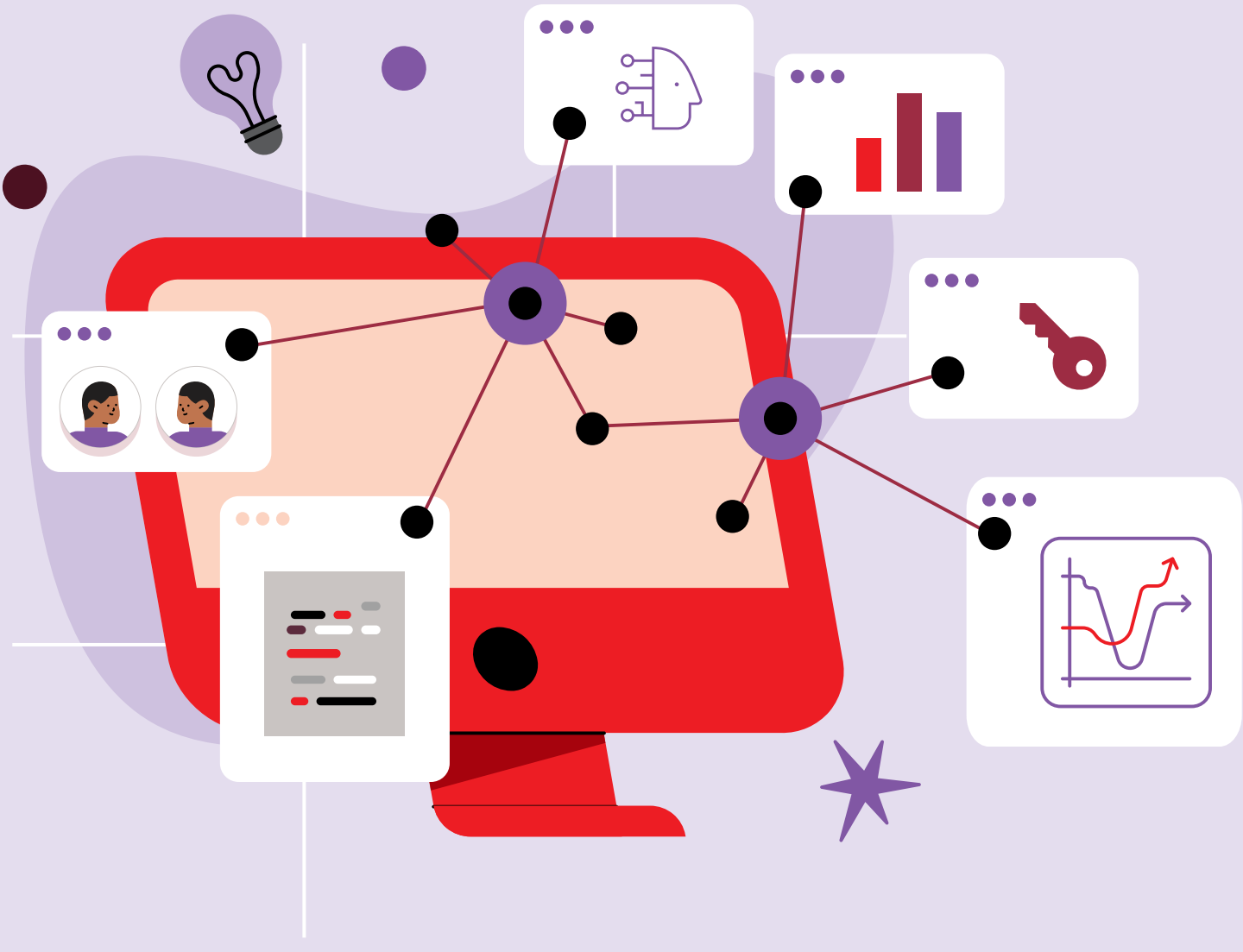
International standards development processes require inclusive participation from diverse stakeholders, such as SMEs, consumers, civil society, and underrepresented communities, to ensure that AI-related standards are effective and responsive to varied needs market and societal interests.

8

Conformity assessment, underpinned by international standards, enables robust verification of AI systems' compliance with policies, regulations and technical standards, fostering trust in and wide adoption of AI technologies.

6. Recommendations

To ensure that AI governance frameworks are inclusive, effective, and globally aligned, coordinated action is needed across key stakeholder groups. The following recommendations outline targeted roles and responsibilities for policymakers, standards bodies, industry stakeholders, and civil society in advancing trustworthy and interoperable AI through standardization.



6.1 For policymakers

- Integrate standardization expertise into the life cycle of policy development by actively engaging with national standards bodies during AI strategy formulation, implementation and evaluation.
- Participate in international standards development to ensure regulatory needs are addressed and global interoperability is enhanced.
- Consider referencing international standards in regulatory frameworks to promote interoperability, reduce compliance costs, and facilitate cross-border AI deployment.

6.2 For standards bodies

- Continue developing AI-specific standards that consider both technical and socio-technical dimensions.
- Enhance inclusive stakeholder engagement for AI-related standards to ensure standards reflect diverse perspectives and varied real-world needs and implementation challenges.
- Strengthen coordination mechanisms with local governments, as well as regional and international policy bodies, to ensure standards effectively support emerging AI governance frameworks.

6.3 For industry stakeholders

- Actively participate in standards development processes to contribute technical and socio-technical expertise and ensure standards reflect industry best practices.
- Adopt standards-based approaches to AI governance to demonstrate responsible innovation, streamline regulatory compliance, and facilitate global market access.
- Implement conformity assessment mechanisms to verify and communicate standards compliance to customers, regulators, and the public.

6.4 For civil society

- Engage in standards development to ensure diverse societal perspectives, particularly from potentially affected communities, are represented.
- Collaborate to build the capacity of civil society representatives, consumer organizations, and standard development organizations to work together.
- Contribute expertise on societal implications of AI to strengthen the socio-technical foundations of AI standards.
- Monitor standards implementation to identify gaps and help ensure AI governance frameworks address present harms and potential future risks.

7. How can I get involved?

There are multiple pathways for engaging with the development and implementation of AI standards – whether through national standards bodies or as an individual stakeholder.

7.1 NSBs:

To follow or contribute to ISO's technical work on AI and its governance, see the relevant Technical Committees:

- ISO/IEC JTC 1, Information Technology, and its subcommittees:
  - **ISO/IEC JTC 1/SC 42, *Artificial Intelligence***
  - ISO/IEC JTC 1/SC 27, *IT Security Techniques*
  - ISO/IEC JTC 1/SC 32, *Data management and interchange*
  - ISO/IEC JTC 1/SC 38, *Cloud computing and distributed platforms*
  - ISO/IEC JTC 1/SC 40, *IT Service management and IT governance*
  - ISO/IEC JTC 1/SC 41, *Internet of Things and Digital Twin*
  - ISO/IEC JTC1/SC42 and ISO/CASCO Joint Working Group: *Conformity assessment schemes for AI systems*

Experts with sector-specific knowledge, especially regarding the application of AI in particular industries, are invited to contribute to the relevant ISO Technical Committees on the sector, many of which are incorporating AI-related workstreams.

7.2 Stakeholders:

- For more information about this publication and ISO's work on AI policy and governance, please contact the ISO Capacity Building Unit (capacity@iso.org).
- If you have expertise related to AI and its governance and would like to contribute to the development of standards, please contact the ISO member in your country (www.iso.org/members.html).
- For more information about ISO standards for AI and its governance, we encourage you to contact the secretariat of ISO/IEC JTC 1/SC 42, Artificial Intelligence (contact details available at: www.iso.org/committee/6794475.html).

8. References

1. African Union. (2024). Continental strategy on artificial intelligence. https://au.int/sites/default/files/documents/44004-doc-EN-_Continental_AI_Strategy_July_2024.pdf
2. Association of Southeast Asian Nations (ASEAN). (2024). ASEAN guide on AI governance and ethics. https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics_beautified_201223_v2.pdf
3. Chen, B.J., & Metcalf, J. (2024). Explainer: A Sociotechnical Approach to AI Policy. Data & Society. https://datasociety.net/library/a-sociotechnical-approach-to-ai-policy/
4. China, State Council of the People's Republic of China. (2017). A new generation of artificial intelligence development plan. https://flia.org/notice-state-council-issuing-new-generation-artificial-intelligence-development-plan/
5. Council of Europe. (2024). Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law. https://rm.coe.int/1680afae3c
6. European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council on artificial intelligence (AI Act). Official Journal of the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689
7. Government of Canada. (2021). Canada's artificial intelligence strategy. https://ised-isde.canada.ca/site/ai-strategy/en
8. International Organization for Standardization (ISO). (2023). Building trust: The conformity assessment toolbox. https://www.iso.org/publication/PUB100230.html
9. Jonker, A., & Gomstyn, A. (2024). What is AI alignment? IBM. https://www.ibm.com/think/topics/ai-alignment
10. Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., & Clark, J. (2024). The AI Index 2024 Annual Report. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. https://hai.stanford.edu/ai-index/2024-ai-index-report
11. Ministry of Communications and Information Technology (MCIT), Egypt. (2025). Egypt National Artificial Intelligence Strategy, 2025-2030, 2nd Edition. https://ai.gov.eg/SynchedFiles/en/Resources/AIstrategy%20English%2016-1-2025-1.pdf
12. Ministry of Economy, Trade and Industry (METI), Japan. (2022). AI Governance in Japan Ver. 1.1. https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20210709_8.pdf
13. Ministry of Information, Communications and the Digital Economy of Kenya. (2025). Kenya Artificial Intelligence Strategy 2025-2030. https://www.ict.go.ke/sites/default/files/2025-03/Kenya%20AI%20Strategy%202025%20-%202030.pdf
14. Ministry of Science, Technology and Innovation (MCTI), Brazil (2021). Summary of the Brazilian Artificial Intelligence Strategy. https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia-summary_brazilian_4-979_2021.pdf
15. National Centre for Artificial Intelligence and Robotics (NCAIR), Nigeria. (2024). Nigeria National Artificial Intelligence Strategy (DRAFT). https://ncair.nitda.gov.ng/wp-content/uploads/2024/08/National-AI-Strategy_01082024-copy.pdf
16. National Institute of Standards and Technology (NIST). (2023). NIST AI risk management framework. https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf

17. Organisation for Economic Co-operation and Development (OECD). (2024). OECD AI Principles. OECD. https://oecd.ai/en/ai-principles

18. Organisation for Economic Co-operation and Development (OECD). (2024). Recommendation of the Council on Artificial Intelligence. OECD. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

19. Sætra, H.S., & Danaher, J. (2025). Resolving the battle of short- vs. long-term AI risks. AI Ethics 5, 723–728. https://doi.org/10.1007/s43681-023-00336-y

20. State Council of the People's Republic of China. (2017). A new generation of artificial intelligence development plan. https://flia.org/notice-state-council-issuing-new-generation-artificial-intelligence-development-plan/

21. The White House, Office of Science and Technology Policy. (2022). The AI Bill of Rights: Protecting civil rights in the age of artificial intelligence. The White House. https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/

22. United Nations. (2024). United Nations Global Digital Compact: A shared vision for the digital future. https://www.un.org/en/global-digital-compact

23. United Nations Educational, Scientific and Cultural Organization (UNESCO). (2021). Recommendation on the ethics of artificial intelligence. UNESCO. https://en.unesco.org/artificial-intelligence/ethics

24. United States Congress. (2020). H.R.6216 - National Artificial Intelligence Initiative Act of 2020. https://www.congress.gov/bill/116th-congress/house-bill/6216

25. UK Government. (2021). National AI strategy. https://www.gov.uk/government/publications/national-ai-strategy

26. UK Government. (2023). AI regulation: A pro-innovation approach [White paper]. https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper

About ISO

ISO (International Organization for Standardization) is an independent, non-governmental international organization with a membership of 173* national standards bodies. Through its members, it brings together experts to share knowledge and develop voluntary, consensus-based, market-relevant International Standards that support innovation and provide solutions to global challenges.

ISO has published more than 25 900* International Standards and related documents covering almost every industry, from technology to food safety, to agriculture and healthcare.

For more information, please visit **www.iso.org**.

*August 2025

ISO website: www.iso.org
ISO newsroom: www.iso.org/news
ISO videos: www.iso.org/youtube
Follow us on Twitter: www.iso.org/twitter
Join us on Facebook: www.iso.org/facebook

International Organization for Standardization

ISO Central Secretariat
Chemin de Blandonnet 8
1214 Geneva, Switzerland

Discover the digital version !

iso.org

We care about our planet.
This brochure is printed on recycled paper.
© ISO, 2025
All rights reserved
ISBN 978-92-67-11461-3