

*Commission on Science and Technology for
Development (CSTD)*

*Developing International
Governance Mechanisms to
Regulate the Cross-Border
Impacts and Assign Legal
Liability for Artificial
Intelligence Systems,
Including Challenges of
Enforcement, Algorithmic
Accountability, and Alignment
of International Standards.*

*Director: Ana Sofia Arriaga Chan
Moderator: Juliette Camacho Albert*

I. INTRODUCTION

Considered one of the principal organs of the United Nations and a subsidiary body of the Economic and Social Council (ECOSOC), the United Nations Commission on Science and Technology for Development (CSTD) was founded in 1992 (Commission on Science and Technology for Development | UNCTAD, n.d.). This was achieved through General Assembly Resolution 46/235, which revitalized and restructured the United Nations' economic and social fields, thereby making the committee a functional commission of ECOSOC (Mandate and Institutional Background | UNCTAD, n.d.). Its sessions are held in April 1993 in New York, with the Palais des Nations in Geneva, Switzerland, as the main places where the United Nations Conference on Trade and Development (UNCTAD) has had the purpose to function as a secretariat since July 1993 (Mandate and Institutional Background | UNCTAD, n.d.).

Around the time it was established, technology and science were taken as important factors that hugely influenced global development, where technological developments defined the benefits for all nations, especially in developing countries (About the CSTD | UNCTAD, n.d.). Looking back, the 1979 UN Conference on Science and Technology for Development, held in Vienna initially, established an intergovernmental committee, whereas passed to be a functional commission of ECOSOC in 1992; the General Assembly elevated the status and mandate of the body, showing the rapid growth of technological change after the Cold War (About the CSTD | UNCTAD, n.d.). The CSTD has around 43 member states elected by ECOSOC, with four-year terms and membership allowed only to national governments; civil society organizations are fundamental for this organization's deliberations (UN Commission on Science and Technology for Development (CSTD), 2024).

The core mandate is usually directed to the United Nations General Assembly and ECOSOC, where its science and technological issues influence and contribute to the policy applications derived by those two committees (Yinuo & Yinuo, 2026). Therefore, it is considered an intergovernmental forum where countries can comment on issues and opportunities that technological development has impacted, ensuring that developing countries are also actively involved (Yinuo & Yinuo, 2026). ECOSOC has mandated this committee since 2005, working as a focal point for the system follow-up outcomes of the World Summit on the Information Society (WSIS). With the recent adoption of the Global

Digital Compact in September 2024, the CSTD worked on Group Data governance, showing an expanding role in shaping the digital outcomes (Working Group on Data Governance at All Levels, 2023).

Significant limitations in the committee represent a constant challenge in the force that it has to influence in other bodies because it is an advisory body; the imposition of obligations of recommendations can't be applied to member states, as this can only be taken by ECOSOC and the General Assembly when drafting solutions (Economic and Social Affairs, Economic and Development Issues, Development Bodies and Programs, Commission on Science and Technology for Development, n.d.). CSTD needs to maintain a stable political will and how states choose to implement the guidance that is concluded by CSTD, defining the effectiveness of the work (Economic and Social Affairs, Economic and Development Issues, Development Bodies and Programs, Commission on Science and Technology for Development, n.d.). Furthermore, the consensus-based process established that positions could be diluted to adapt into the broadest possible agreement. These limitations had been a challenge for the committee, but important recommendations made had contributed to important resolutions applied, such as shaping international norms around technology and development; addressing digital divisions in the 2000s; and exploring the governance frontier of technologies, recently, including Artificial Intelligence (AI) (The Need for an Honest Discussion on Technology and Development, 2019).

Regarding the topic being discussed, the CSTD core focus is to explore the governance challenges of the 21st century: Developing International Governance Mechanisms to Regulate Cross-Border Impacts and Assign Legal Liability for Artificial Intelligence Systems. Meaning that AI has the potential to be a priority tool that will reshape existing economies, healthcare, education, criminal justice, warfare, and virtually all systems (CSTD Side Event: Artificial Intelligence for Regional Sustainable Development, 2024). The implementation of technology has to be considered as something that has to be monitored, with a clear legal framework that regulates it and has accountability, as, on the contrary, it can cause social harm (CSTD Side Event: Artificial Intelligence for Regional Sustainable Development, 2024).

Therefore, it is important to discuss and draft how AI can be effective in serving legal systems, economic contexts, and geopolitical interests. It must be ensured that technology

does not reinforce existing inequalities between the North and the South, allowing impunity and undermining the rule of law, are the principal issues that limit the implementation of these technological tools (28th CSTD Side Event: Inclusive Artificial Intelligence for Development, 2023). An efficient, structured international governance architecture can use AI for enhancing sustainable development, protecting fundamental rights, and creating a more equitable world. The CSTD, following the core mandate for science, technology, and development, is the precise body needed to draft meaningful progress (28th CSTD Side Event: Inclusive Artificial Intelligence for Development, 2023).



II. HISTORICAL CONTEXT

The Birth of AI and Early Governance Debates (1950s-2000s)

Alan Turing and John McCarthy were among the pioneers of artificial intelligence development in the 1950s, although its real-world applications remained limited, since questions of governance were absent from the international agenda (Elamin, 2024). Until the 1980s and 1990s, formal discussion on AI governance began as expert systems started to appear in the commercial market, and later in the 2000s with the rise of machine learning (Anastasiia, 2025). Artificial Intelligence initially was treated similarly to telecommunications; it was not recognized as a distinct governance challenge. Meaning, that AI lacked governance frameworks, where intelligent systems were standard software or basic technology (Anastasiia, 2025).

Until the 2000s, international discussions began to take place to discuss the governance of emerging digital technologies (UNCTAD, 2025). The World Summit on the Information Society (WSIS), held in Geneva in 2003 and Tunis in 2005, was a United Nations organization that had the purpose of bridging the digital divide and ensuring that global information and communication technologies were used to foster sustainable development. With these principles stated, the WSIS produced landmark declarations and action plans to bridge the divide to protect human rights online, and promote International cooperation in information and communication technologies (ICTs) (World Summit on the Information Society, 2003). The summit declared international digital governance frameworks applied to intelligent technologies at the time but not specifically to AI since it was not relevant at the moment, but they were later applied to debate the regulations of them (World Summit on the Information Society, 2003).

The Machine Learning Revolution and Rising Concerns (2010-2016)

While in the 1950s the foundational questions towards machine intelligence were stated, researchers developed a clearer understanding of its capabilities in the 2000s (Mijares, 2025). Around 2012, deep learning techniques improved in the sense that AI could perform image recognition tasks. With this demonstration, significant massive investments and research were

done for developing AI systems capable of outperforming humans in specific tasks, such as medical imaging, language translation, and legal document review (Mijares, 2025).

Throughout that decade, the use of artificial intelligence began to increase, leading to the first major documentation of cases in which AI systems caused harm (Lavanchy, 2022). The first case was documented at Amazon, as the business started to use an AI-driven recruitment tool that systematically discriminated against women since it was programmed with male-dominated historical hiring data (Lavanchy, 2022). Then, the agency of ProPublica initiated an investigation in 2016 of the COMPAS AI tool used in U.S. courts, where it was found that the algorithm was likely to falsely flag Black defendants as future criminals compared to white defendants, demonstrating that the algorithmic bias could produce discriminatory outcomes in high-stakes judicial decisions (Module 2: Real-Life Examples of Bias | OLCreat, 2016). Therefore, the existence of these reports was raising concern regarding the fact that the AI systems were making consequential decisions in individuals lacking legal liability for errors and not allowing transparency (Module 2: Real-Life Examples of Bias | OLCreat, 2016).

Consequently, across different countries, this was starting to be a cross-border international problem, since similar cases continue. In the United States, a company deployed an artificial intelligence system that could be deployed via a cloud service in Brazil; it was able to analyze data from users in Nigeria and made decisions that affected individuals in the United Kingdom (Council of Europe, 2024). These problems began to be a crucial matter for legal scholars and policymakers, with the need to create a legal framework that addresses this international issue (Council of Europe, 2024).

National Frameworks Emerge Amid Growing Calls for International Standards (2017-2021).

Between the periods 2017 and 2021, a proliferation of national AI strategies and ethical guidelines started to emerge without a coordinated international framework that managed their effectiveness (Prosper, 2025). Starting with China, it adopted its New Generation Artificial Intelligence Development Plan to become the world's leading AI power by 2030 and transform AI into a potential strategy that needed national priority (Prosper, 2025).

Furthermore, the European Union established a set of regulatory frameworks to manage the AI ethics framework that will be known in several years as the EU AI Act (European Parliament, 2025). Under the Obama and Trump administrations, the United States declared a series of AI strategy documents and avoided applying comprehensive regulation, as it affected the support that innovation needed (European Parliament, 2025).

Due to the increasing matter of regulating AI use, the Organization for Economic Co-operation and Development (OECD) adopted principles on artificial intelligence that established the first intergovernmental AI policy used across 40 countries, including India, Brazil, and several African nations (EvalCommunity Academy, 2025a). It addressed the need for AI systems to be designed as inclusive, transparent, accountable, and safe technology (EvalCommunity Academy, 2025a). Still, several limitations existed, since a limited group of high-income countries was not bound; their practical impact depended on voluntary domestic implementation (EvalCommunity Academy, 2025a). Following important organizations, UNESCO developed its Recommendation on the Ethics of AI, formally adopted in November 2021 by the 193 UNESCO member states (EvalCommunity Academy, 2025). This contribution addressed AI ethics, including human rights, transparency, fairness, privacy, and the governance of AI, since the international community started to recognize AI governance as a shared global responsibility (EvalCommunity Academy, 2025b).

Furthermore, the use of AI in social media began to increase significantly, demonstrating on a global scale the potential harms that an unregulated AI algorithm could inflict on decision-making (Ali et al., 2024). Therefore, content recommendation algorithms started to be optimized for encouraging the spread of political disinformation, incitement of ethnic violence in Myanmar in 2017, and mass radicalization (Ali et al., 2024). In 2021, Meta conducted internal research, and it confirmed the serious harms from its algorithms (Pazzanese, 2026). Questions regarding legal responsibility also emerged in which someone had to be attributed, started to raise questions and being a matter of continuous debate because society started to question who should be held responsible for AI-related harm. Was the responsibility of the developer, the deployer, the platform, or the state, which remained unresolved (Pazzanese, 2026).

The Generative AI explosion and urgent global governance demands (2022-2023).

The release of ChatGPT in November 2022 intensified these social issues (Mesko, 2023). The program, developed by the OpenAI company, was increasingly being used by society, which demonstrated the broad capabilities and potential risks of generative AI, such as generating misinformation, contributing to concerns about job displacement, enabling academic fraud, and causing privacy violations, and the potential that it had to be used as autonomous weapon systems (Mesko, 2023). These AI system intensified market competition, where similar models were developed by companies, generating alternatives for ChatGPT use, such as Google, Meta, Anthropic, and Baidu (Mesko, 2023).

Further, governments started to take action but with different purposes, creating a fragmented regulatory landscape that continues to shape international AI governance (Perboli et al., 2025). For example, the European Union continued work towards implementing the EU AI Act, reaching a political agreement in December 2023 (Perboli et al., 2025). Then China implemented measures for the management of Generative Artificial Intelligence Services in August 2023, with an ideological alignment and national security focus (Perboli et al., 2025). The United States under the Biden administration issued an Executive Order on AI safety in October 2023, establishing an AI safety institute that called for federal agencies to develop AI risk management measures (Perboli et al., 2025).

This situation was reached through the United Nations Security Council, which held an unprecedented session regarding the AI's implications for peace and security in 2023, which marked the first time the Security Council formally discussed artificial intelligence and to discuss Artificial Intelligence (Security Council Report, 2023). This session served to discuss specific concerns, including autonomous weapons, AI-enabled cyberattacks, and disinformation, showing the shift of AI governance from the domain of technology policy to the center of international security (Security Council Report, 2023).

Subsequently, in 2023, the Bletchley Declaration of November, signed by 28 countries, including the United States, the United Kingdom, China, and the European Union, enacted the first major multi-stakeholder agreement on AI safety at an international summit (Burgess Salmon LLP, 2023). Importantly, the declaration was non-binding and focused on frontier AI safety, without addressing the broader questions of cross-border liability,

algorithmic accountability, and inclusive governance that are crucial for the committee to address (Borges Salmon LLP, 2023).



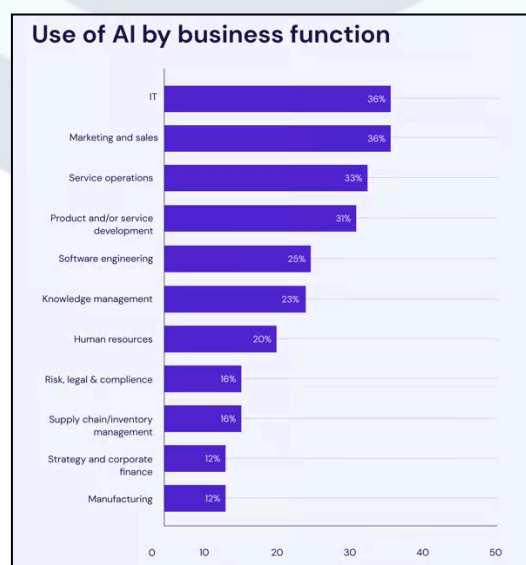
III. CURRENT SITUATION:

International AI governance has faced several challenges within the period from 2024-2026, including developments where the EU AI Act is active, the first UNGA resolution on AI, and the establishment of new UN governance strategies to give it continuation within the urgency of the problem (United Nations, n.d.). At the same time, these developments impose fundamental tensions in the committee.

The Regulatory Fragmentation Crisis

The world's first comprehensive horizontal AI regulation was enacted by the European Union, known as the “AI Act”, in 2024 (European Commission, n.d.). This Act categorizes AI systems as prohibited in certain applications, including social scoring by governments, healthcare, education, and law enforcement (European Commission, n.d.). This Act's measures are being applied from 2024 to 2026, with prohibitions such as categorizing them on unacceptable risk applications applying from 2025, and most high-risk provisions are expected to be applied in 2026 (Decode the Future, n.d.). This Act has become a factual global benchmark where companies that are used and recognized on a worldwide scale are obligated to comply with the requirements for having access to the European Market (Decode the Future, n.d.).

Figure I.



Source:

The growing adoption of AI across important functions in worldwide-use companies that have legal liability frameworks is continually increasing (Hostinger, n.d.). The primary operational functions identified as requiring strict human oversight, with 20% of AI-dependent use; service operations, with 33%; information technology, with 36% (Hostinger, n.d.). As a result, it explains the discriminatory results, privacy violations, and inaccurate decisions that affect a worldwide scale across national borders (Hostinger, n.d.).

The United States can be considered the one taking a different approach; in 2025, with the change in administration from Biden, the AI safety framework was restored, and the U.S. favored sector-specific regulatory measures (VerifyWise, 2026). In 2026, approximately 1,000 AI-related bills were included in the U.S. state legislatures in 2025 alone, which created certain tensions with state-level requirements (VerifyWise, 2026). Whereas China continues to enforce AI regulations by proposing its own International AI governance organization (Institute of Public Administration and Governance, n.d.). Therefore, three major AI powers with incompatible frameworks are creating challenges to resolve conflicts within cross-border AI deployment and demonstrate the lack of effective international coordination decisions (Institute of Public Administration and Governance, n.d.).

The United Nations Conference on Trade and Development (UNCTAD) has documented that one-third of developing countries have implemented national AI strategies, and approximately 118 developing nations remain excluded from global AI governance, creating a significant differentiation (International Trade and Compliance Center, n.d.). For example, Africa has less than 1% of global data capacity, requiring \$2.6 trillion in investment by 2030 to address the bridge in the digital infrastructure gap (International Trade and Compliance Center, n.d.).

The Legal Liability Gap and Documented Harms

The absence of legal liability frameworks for AI-caused harm remains an unresolved issue that must be addressed; for example, traditional product liability law was applied to conclude with probabilistic, emergent behavior that caused harm since it is difficult to attribute it to a specific programming decision (Sharma & Mishra, 2024). Where this AI systems are applied

in healthcare and misidentify a patient's cancer (Sharma & Mishra, 2024), or when a resume screening algorithm systematically excludes qualified candidates based on protected characteristics and also when autonomous vehicles fail to recognize pedestrians (Seyfarth Shaw LLP, 2024), the legal responsibility remains null since there is not legal responsibility directed to a specific person (Sharma & Mishra, 2024).

A highlighted AI issue was the case of *Mobley V. Workday Inc.*, where AI systems conducted recruitment employment discrimination, challenging courts in how anti-discrimination law could be applied to hiring systems (Seyfarth Shaw LLP, 2024). This raised concerns with the U.S. Equal Employment Opportunity Commission (EEOC), which imposed enforcement against AI-based tools, where New York City also aligned with placing enforcement against AI-based requirements applied in systems. Regarding Europe, its courts have interpreted the GDPR's provisions in an automated decision-making AI context, being national responses to a fundamental cross-border challenge (GDPR Advisor, n.d.). This cross-border liability paradox is exemplified when an AI system developed in one jurisdiction is deployed by an enterprise in another to evaluate citizens of a third (GDPR Advisor, n.d.).

Therefore, there is potential harm that hasn't been discussed, where AI surveillance systems are approximately deployed by governments in about 75 countries, with documented cases of misidentification with facial recognition that lead to wrong arrests, with various cases being reported in the United States (Voice of America, 2021). Also, historical financial inequalities have been reported as an AI-based error due to scoring systems used globally (Ahmed, 2025). Furthermore, content recommendation algorithms have been documented as driving political polarization and facilitating digital violence (Ahmed, 2025). Attributing to the healthcare sector, AI diagnostic tools and trained data from the West demonstrate the weakness of accuracy when these AI systems are applied and affect demographic groups, and represent a critical concern in developing nations (Ahmed, 2025).

Competing Perspectives and the Global Governance Debate

The current issue is not seen as a technical challenge but has escalated to be recognized as a geopolitical and philosophical division issue by discussing what measures should be necessary to govern AI (Le Diplomate, 2025). It is argued that strong, international standards with strong mechanisms and using the EU model that balances innovation with the protection

of fundamental rights should be the strategies that companies must engage with, deploying the most risky AI applications in jurisdictions with weak control (Le Diplomate, 2025).

Opposed to the European position, the United States and parts of China see national sovereignty as a priority over technology governance, as they are planning to bind international frameworks that could constrain AI industries or intelligence development agencies (Congressional Research Service, 2025). Where premature regulation risks disrupting innovation and international standards that should be voluntary guidelines that allow countries to adapt to modern technology (Congressional Research Service, 2025).

Still, a third perspective is highlighted by the Global South, arguing that both the EU and the United States/China perspectives are discussing regulations without including developing countries, making any governance framework developed for major AI powers risk excluding priorities and values from the vulnerabilities of the majority of the world's population (Center for Strategic and International Studies, 2025a). As the Center for Strategic and International Studies documents, most of the governance regulations that are being developed in Washington, Brussels, and Beijing do not represent the use of AI tools and access from vulnerable countries (Center for Strategic and International Studies, 2025b).

Consequently, in 2025, the United Nations General Assembly enacted (A/RES/79/325), which established an Independent International Scientific Panel on AI and Global Dialogue on AI Governance, but its effectiveness depends on the political will of member states to follow through with substantive action (United Nations, n.d.). Where the intellectual and policy architecture has to be prioritized when discussing regulations, making the global community consider dialogue between all affected countries (United Nations, n.d.).

IV. UN ACTIONS

A series of substantive actions on AI governance have been increasingly taken by the United Nations and its subsequent bodies, reflecting a broad, growing consensus that the issue currently discussed demands a collective multilateral response.

UNESCO Recommendation on the Ethics of AI (2021)

In 2021, UNESCO adopted the Recommendation on the Ethics of Artificial Intelligence, which establishes the universal standard for ethical AI. Its execution has led about 75 member states to leverage this framework to have inclusive, human-centered AI governance, addressing risks such as bias, surveillance, and environmental degradation (United Nations Educational, Scientific, and Cultural Organization, n.d.). This adoption encourages member states to build national policies, regulatory frameworks, and AI ethics mechanisms to provide a governance framework that includes the UNESCO Readiness Assessment Methodology (RAM) that supports countries in the evaluation of the capacity that they hold to implement ethical AI standards (United Nations Educational, Scientific and Cultural Organization, n.d.).

UNGA Resolution on Safe, Secure and Trustworthy AI (March 2024)

The Resolution A/78/L.49 or A/RES/78/265 (2024) represents the first global consensus effort to establish governance for artificial intelligence, adopted unanimously by all 192 member states (United Nations General Assembly, 2024). The United States officially spearheaded and endorsed the initiative, with 123 member states cosponsoring it, emphasizing human rights, bridging the digital divide, and enabling developing countries to engage effectively by boosting their ability to utilize AI (United Nations General Assembly, 2024). This effort is further supported by the EU AI Act and U.S. Executive Orders that are influencing the worldwide governance of AI. Thus, the UN resolution stresses technical and financial support to developing countries to bridge the technology divide, and the EU AI Act mandates that AI systems be created in a way that upholds basic liberties, confidentiality, and avoidance of discrimination (Government of the United Kingdom, 2023). Regarding cross-border cooperation, the U.S. government led the resolution to establish a framework that will influence international frameworks like the EU Artificial Intelligence Act and the

Bletchley Declaration. The White House Executive Order requirements for testing and risk management protocols for applying governance standards worldwide (Government of the United Kingdom, 2023).

Global Digital Compact and High-Level Advisory Body Final Report (2024)

In September 2024, at the “Summit of the Future” and as part of the Pact for the Future, the United Nations adopted the Global Digital Compact (GDC)—a landmark agreement on the governance of digital technologies, including AI. The GDC committed to closing the digital divide, regulating emerging technologies like AI according to human rights, and guaranteeing that developing countries have a real seat at the table in shaping digital norms (United Nations, n.d.). However, in the same year, the final report of the UN Secretary-General's High-Level Advisory Body on Artificial Intelligence was released under the title “Governing AI for Humanity” and provided a detailed framework for an international regime of AI governance (United Nations, n.d.). The report proposed establishing a new international regime of AI governance (United Nations, 2024). The report proposed establishing a new international panel of scientists on AI that would be similar to the IPCC, an annual global policy dialogue on AI governance, and a specific capacity-building initiative for developing nations (United Nations, 2024).

UNGA Resolution Establishing International Scientific Panel and Global Dialogue on AI (August 2025)

On 26 August 2025, the General Assembly adopted resolution A/RES/79/325 establishing two important AI governance mechanisms (United Nations General Assembly, 2025). First, it created an Independent International Scientific Panel on Artificial Intelligence (United Nations General Assembly, 2025). It is multidisciplinary, with 40 members who serve three-year terms, selected through a transparent process to ensure balanced geographic and gender representation (United Nations General Assembly, 2025). The Panel is mandated to produce annual, evidence-based scientific assessments of AI relevant to policy and to operate according to principles of independence, credibility, and inclusivity (United Nations General Assembly, 2025).

Second, the resolution launched a Global Dialogue on Artificial Intelligence Governance, an annual forum for world leaders and international stakeholders to discuss critical issues of AI governance, including safety, ethics, rights, transparency, capacity-building, interoperability, and closing the digital divides (United Nations, n.d.). These mechanisms are the most significant institutional step so far toward a coherent international AI governance architecture, but their success will ultimately depend on sustained political will and adequate resourcing (United Nations, n.d.).

ECOSOC Special Meeting on AI (May 2025)

In May 2025, ECOSOC held a Special Meeting on artificial intelligence devoted specifically to the subject, as it was becoming more prominent on the UN development agenda (President of the Economic and Social Council, 2025). The gathering underscored the pressing importance of addressing digital and data divides, noting that 2.6 billion people remain unconnected and that the bulk of AI development is taking place without sufficient input from the world's most vulnerable populations (President of the Economic and Social Council, 2025).

Key recommendations from the meeting included addressing the digital infrastructure gaps, especially in the Global South; ensuring equitable access to computing resources; and translating the principles of the Global Digital Compact into coordinated action (United Nations Conference on Trade and Development, 2025). The meeting also underscored the important role of the CSTD in AI governance due to its unique mandate linking science, technology, and development (United Nations Conference on Trade and Development, 2025).

V. POSSIBLE SOLUTIONS

The problems highlighted by this background paper are fragmented national systems, a lack of transborder liability systems, a lack of algorithmic accountability, jurisdictional disputes, and the exclusion of developing countries, which are fundamentally interrelated (United Nations, 2024; OECD, 2024). Any successful approach must take into consideration all these aspects together in a manner that is internationally legitimate, practical in enforcement, and developed with input from both developed and developing countries.

I. Institute Mandatory Algorithmic Impact Assessments for Cross-Border AI Deployments.

Due to the rapid growth of border-crossing AI technology, there are growing concerns about algorithmic biases, lack of transparency, and accountability in areas like healthcare, education, financial services, the labor market, and policing (OECD, 2024). Due to significant differences between laws in countries, the lack of a uniform method of risk assessment hampers governments' capacity to analyze the possible consequences of the use of the systems (UNESCO, 2021). Thus, CSTD should facilitate an international framework of mandatory AIAs (United Nations, 2024).

- a) Short term: Encourage the application of AI systems that are deployed internationally to conduct Algorithmic Impact Assessments (AIAs) before deployment.
- b) Medium term: Establish an international publicly available registry of AIAs, enabling government officials and other parties to analyze the transparency and risk assessments performed by those algorithms.
- c) Long term: Implement cooperative enforcement systems whereby exporting countries will certify AIAs but importing countries will have the power to monitor or even suspend use of high-risk systems.

II. Create an International Standards Harmonization Body for Artificial Intelligence

The rapid proliferation of fragmented national AI regulations has generated severe standardization inconsistencies, complicating cross-border deployment, which makes the implementation of AI across borders difficult and poses further problems for developing

countries that lack regulatory capacity (International Organization for Standardization & International Electrotechnical Commission [ISO/IEC], n.d.; International Telecommunication Union, n.d.). CSTD needs to support the creation of a body that coordinates the international standards for AI (United Nations Conference on Trade and Development, 2024).

- a) Short term: Immediate measures: Establish an International Body on AI Standard Harmonization through cooperation between ITU, ISO/IEC, and UNCTAD.
- b) Medium term: Develop International standards on risk assessment, audit, and conformity testing for AI that are mutually recognized, as well as support participation of developing countries.
- c) Long term: Facilitate a global mutual recognition regime for AI systems that have been certified by participating standards and operate in member states, except in cases of national security.

III. *Establish a Global AI Capacity Building and Technology Transfer Fund*

The uneven ability of states to develop and govern AI technologies will restrict the participation of developing states in the creation of international standards for AI governance (United Nations Conference on Trade and Development, 2024). Without adequate skills, funding, and technology, it will be challenging for many states to adopt responsible governance of AI technologies (United Nations, 2024). Hence, CSTD should encourage international cooperation through global capacity building and technology transfer (United Nations Conference on Trade and Development, 2024).

- a) Short term: Establish a Global AI Capacity Building and Technology Transfer Fund within the United Nations Conference on Trade and Development (UNCTAD), funded by the world's leading producers of AI technology.
- b) Medium term: Fund training programs for regulatory, policymaking, and judicial bodies, as well as fund national research and computing infrastructure and open data projects.

- c) Long term: Facilitate technology transfer that allows developing nations to develop AI technology suited to their own national interests.

IV. Establish a Binding International Framework Convention on AI Governance.

The lack of an international agreement on the regulations of AI will lead to inconsistencies in its governance, insufficient accountability, and a lack of clarity in terms of liability in case of cross-border disputes (United Nations, 2024). Member countries need to develop an international convention on the regulation of AI (United Nations Conference on Trade and Development, 2024).

- a) Short term: Discuss the UN Framework Convention that sets out the minimal standards for the safe use of AI.
- b) Medium term: Implement mandatory incident reporting, implementation plans, and an international secretariat that oversees the process of review.
- c) Long term: Implement a liability protocol that sets out jurisdictional standards and international claims procedures, where the AI developers and deployers are responsible for any damage caused by AI.

V. Develop an International Mechanism for Inclusive AI Governance Participation.

Current deliberations about AI governance continue to be dominated by technologically advanced nations, while some developing nations have no representation in international governance, even though they are greatly affected by the adoption of AI (United Nations, 2024). The CSTD needs to advocate for the creation of an inclusive AI governance structure (United Nations Conference on Trade and Development, 2024).

- a) Short term: Create a permanent UN Intergovernmental Panel on AI Governance that should be organized under ECOSOC for coordination of international cooperation.
- b) Medium term: Implement a scientific assessment approach with periodic publications on AI consequences and governance to assist in the negotiation process regarding international AI standards.

Bind the participation of developing countries in capacity building, participation in decision-making, and more influence in the international consensus-building process.



VI. COUNTRIES INVOLVED

1. *United States of America*

The United States is known to support a voluntary and innovation-led approach for governance of AI technologies, such as its lead role in crafting the UN General Assembly's first AI resolution, which was adopted by consensus on March 21, 2024, and co-sponsored by 123 member states of the UNGA, including China (United Nations General Assembly, 2024). At the domestic level, the establishment of the AI Safety Institute under NIST is part of efforts by the United States to facilitate testing and evaluation standardization efforts with industry (National Institute of Standards and Technology, 2024). Despite being a party to the Council of Europe's Framework Convention on AI, signed in September 2024, the United States has remained steadfast in resisting all attempts at imposing any kind of international framework that may limit its competitiveness or affect its intelligence and national security uses of AI Technology (Council of Europe, 2024).

2. *French Republic*

In terms of multilateral policy initiatives, France has positioned itself as Europe's leading AI governance, following up on the EU's AI Act as well as active participation in the negotiation of the Council of Europe's Framework Convention on Artificial Intelligence, the world's first legally binding international instrument regulating AI, opened for signature in Vilnius in September 2024 (European Union, 2024). Having been one of the Council of Europe's founding members, France advocated for provisions obliging states to ensure that AI systems abide by the principles of human rights, democracy, and the rule of law across their life cycle, excluding applications related to national security (Council of Europe, 2024). France has also become the host of the most important international AI diplomacy initiatives, such as the 2025 AI Action Summit held in Paris (Government of France, 2025).

3. *Russian Federation*

The Russian Federation has an emphasis on national sovereignty and technological independence, which is evident from its National AI Strategy to 2030, which has been extensively revised by Presidential Decree No. 124 in February 2024 to respond to sanctions

imposed by the West following the 2022 invasion of Ukraine (President of the Russian Federation, 2024). According to Russia's submission to CSTD in 2024 and supported by more than 390 organizations, it insists that international cooperation should respect the right of each state to create its own regulatory framework for AI (United Nations Conference on Trade and Development, 2024). Due to sanctions limiting access to sophisticated semiconductors, Russia is increasingly cooperating with China and India regarding AI and refusing any oversight by the West (Carnegie Endowment for International Peace, 2024).

4. *People's Republic of China*

The People's Republic of China has taken a more aggressive stance toward global AI governance in relation to the 2023 Global AI Governance Initiative put forward by President Xi Jinping (State Council of the People's Republic of China, 2023). In July 2025, Premier Li Qiang unveiled the Global AI Governance Action Plan, offering to set up the World AI Cooperation Organization based in Shanghai to work alongside the United Nations in resolving the “AI divide” among Global South Nations (State Council of the People's Republic of China, 2023). China, on one hand, is working to enact its own set of comprehensive domestic policies on generative AI, and, on the other hand, pursues the course of multilateralism while rejecting any Western attempts at imposing their standards, claiming that they only serve to maintain technological hegemony (Xinhua, 2025). China co-sponsored the 2024 UN General Assembly resolution on AI with the United States (United Nations General Assembly, 2024).

5. *The United Kingdom of Great Britain and Northern Ireland*

The United Kingdom, while maintaining its distinctive and innovative approach towards AI governance based on risk management that is different from but not in contradiction with the EU AI Act, also played a key role in AI safety diplomacy internationally (UK Government, 2023). The first AI Safety Summit was held at Bletchley Park in the United Kingdom in November 2023, resulting in the adoption of the Bletchley Declaration, signed by 28 states, including the United States, China, and the European Union, and the signing of the Council of Europe's Framework Convention on AI in September 2024 (UK Government, 2023). An information exchange partnership in assessing frontier AI models has been established by the

UK's AI Safety Institute with American and other foreign counterparts (Council of Europe, 2024). In the post-Brexit world, the UK's approach combines close cooperation with the EU and US allies and aligns with absolute regulatory independence regarding AI applications in its territory (UK AI Safety Institute, 2024).

6. Federal Republic of Germany

7. State of Japan

8. Republic of Korea

9. Republic of India

10. Dominion of Canada

11. Commonwealth of Australia

12. Italian Republic

13. Kingdom of the Netherlands

14. Swiss Confederation

15. Republic of Singapore

16. State of Israel

17. United Arab Emirates

18. Kingdom of Sweden

19. Federative Republic of Brazil

20. United Mexican States

21. Republic of South Africa

22. Republic of Indonesia

23. Federal Republic of Nigeria

24. Kingdom of Saudi Arabia

25. Republic of Türkiye



VII. BIBLIOGRAPHY

28th CSTD side event: Inclusive Artificial Intelligence for Development. (2023). *UN Trade and Development* (UNCTAD).

<https://unctad.org/meeting/28th-cstd-side-event-inclusive-artificial-intelligence-development>

About the CSTD | UNCTAD. (n.d.). *Unctad.org*.

<https://unctad.org/topic/commission-on-science-and-technology-for-development/about>

Ali, E., Mohamed, S., Osman, M. E., & Balla, B. E. (2024). *The impact of artificial intelligence on social media content. Journal of Social Sciences*.

https://www.researchgate.net/publication/380900493_The_Impact_of_Artificial_Intelligence_on_Social_Media_Content

Anastasiia. (2025). *History of artificial intelligence*. Swiss Cyber Institute.

<https://swisscyberinstitute.com/blog/history-artificial-intelligence/>

Burges Salmon LLP. (2023). *Bletchley Declaration signed at UK AI Safety Summit*.

<https://www.burges-salmon.com/articles/102iryj/bletchley-declaration-signed-at-uk-ai-safety-summit/>

Commission on Science and Technology for Development | UNCTAD. (n.d.). *Unctad.org*.

<https://unctad.org/topic/commission-on-science-and-technology-for-development>

Council of Europe. (2024). *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*.

<https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>

Council of Europe. (2024). *Framework Convention on Artificial Intelligence and human rights, democracy and the rule of law*.

<https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>

CSTD side event: Artificial Intelligence for Regional Sustainable Development. (2024). *UN Trade and Development* (UNCTAD). <https://unctad.org/meeting/cstd-side-event-artificial-intelligence-regional-sustainable-development>

Economic and Social Affairs Economic and Development Issues Development Bodies and Programs Commission on Science and Technology for Development. (n.d.). <https://2009-2017.state.gov/documents/organization/134781.pdf>

Elamin, M. O. I. (2024). *AI through the ages: Unlocking key opportunities and navigating challenges in the history and future of artificial intelligence*. *International Journal of Religion*, 5(12), 1152–1166. <https://doi.org/10.61707/tarzcj89>

European Parliament. (2025). *EU AI Act: First regulation on artificial intelligence*. <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

European Union. (2024). *Artificial Intelligence Act*. <https://artificial-intelligence-act.eu/>

EvalCommunity Academy. (2025a). *OECD AI principles*. <https://academy.evalcommunity.com/oecd-ai-principles/>

EvalCommunity Academy. (2025b). *UNESCO recommendation on AI ethics*. <https://academy.evalcommunity.com/unesco-recommendation-on-ai-ethics/>

Government of France. (2025). *AI Action Summit*. <https://www.elysee.fr/en/ai-action-summit>

Government of the United Kingdom. (2023). *The Bletchley Declaration by countries attending the AI Safety Summit, 1–2 November 2023*. <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>

International Electrotechnical Commission, & International Organization for Standardization. (n.d.). *ISO/IEC JTC 1/SC 42 Artificial Intelligence*. <https://www.iso.org/committee/6794475.html>

International Telecommunication Union. (n.d.). *AI for Good*. <https://aiforgood.itu.int/>

Lavanchy, M. (2022). *Amazon's sexist hiring algorithm could still be better than a human*. IMD.

<https://www.imd.org/research-knowledge/digital/articles/amazons-sexist-hiring-algorithm-could-still-be-better-than-a-human/>

Mandate and Institutional Background | UNCTAD. (n.d.). *Unctad.org*.
<https://unctad.org/topic/commission-on-science-and-technology-for-development/mandate>

Mesko, B. (2023). *The ChatGPT (generative artificial intelligence) revolution has made artificial intelligence approachable for medical professionals*. *Journal of Medical Internet Research*, 25, e48392. <https://doi.org/10.2196/48392>

Mijares, A. L. (2025). *History of AI: Innovation through the decades*. Unity Communications. <https://unity-connect.com/our-resources/blog/history-of-ai/>

National Institute of Standards and Technology. (2024). *U.S. AI Safety Institute*.
<https://www.nist.gov/aisi>

Open University. (2016). *Module 2: Real-life examples of bias*.
<https://www.open.edu/openlearncreate/mod/book/tool/print/index.php?id=227372&chapterid=35267>

Organisation for Economic Co-operation and Development. (2024). *OECD AI Principles*.
<https://www.oecd.org/en/topics/ai-principles.html>

Pazzanese, C. (2026). *Is social media responsible for what happens to users?* *Harvard Gazette*.
<https://news.harvard.edu/gazette/story/2026/02/is-social-media-responsible-for-what-happens-to-users/>

Perboli, G., Simionato, N., & Pratali, S. (2025). *Navigating the AI regulatory landscape: Balancing innovation, ethics, and global governance*. *Economic and Political Studies*, 13(4), 367–397. <https://doi.org/10.1080/20954816.2025.2569584>

- President of the Economic and Social Council. (2025). *Letter to Member States regarding the summary of discussions on artificial intelligence governance*. United Nations Economic and Social Council. https://ecosoc.un.org/sites/default/files/2025-05/250513_PECOSO-letter-for-summary-Member-States.pdf
- President of the Russian Federation. (2024). *National Strategy for the Development of Artificial Intelligence*. <http://kremlin.ru/acts/bank>
- Prosper, D. (2025). *Global standards for AI ethics in education: A comparative review of OECD, UNESCO, and EU guidelines*. https://www.researchgate.net/publication/396563422_Global_Standards_for_AI_Ethics_in_Education_A_Comparative_Review_of_OECD_UNESCO_and_EU_Guidelines
- Security Council Report. (2023). *Artificial intelligence: High-level briefing. What's In Blue*. <https://www.securitycouncilreport.org/whatsinblue/2023/07/artificial-intelligence-briefing.php>
- State Council of the People's Republic of China. (2023). *Global AI Governance Initiative*. <https://english.www.gov.cn/>
- The need for an honest discussion on technology and development. (2019). *UN Trade and Development* (UNCTAD). <https://unctad.org/news/need-honest-discussion-technology-and-development>
- UK AI Safety Institute. (2024). *International collaboration*. <https://www.aisi.gov.uk/>
- UK Government. (2023). *The Bletchley Declaration by countries attending the AI Safety Summit*. <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration>
- UN Commission on Science and Technology for Development (CSTD). (2024). *Civil Society Financing for Development Mechanism*. <https://csforffd.org/un-commission-on-science-and-technology-for-development-cstd>

/

United Nations. (2024). *Governing AI for humanity* [Press release].
https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_press_release.pdf

United Nations. (2024). *Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development* (A/RES/78/265).
<https://documents.un.org/doc/undoc/gen/n24/090/25/pdf/n2409025.pdf>

United Nations. (n.d.-a). *Artificial intelligence and the Global Digital Compact*.
<https://www.un.org/global-digital-compact/en/ai>

United Nations. (n.d.-b). *Global Digital Compact*.
<https://www.un.org/digital-emerging-technologies/global-digital-compact>

United Nations Conference on Trade and Development. (2025). *Summary of discussions: ECN.16/2025*.
https://unctad.org/system/files/information-document/ecn162025_summary-of-discussions_en.pdf

United Nations Conference on Trade and Development. (2025). *Technology and Innovation Report 2025: Inclusive Artificial Intelligence for Development*.
<https://unctad.org/publication/technology-and-innovation-report-2025>

United Nations Conference on Trade and Development. (2025). *Technology foresight and technology assessment for sustainable development* (UNCTAD/TCS/DTL/INF/2025/2). United Nations.
https://unctad.org/system/files/official-document/tcsdtlinf2025d2_en.pdf

United Nations Educational, Scientific and Cultural Organization. (2021). *Recommendation on the Ethics of Artificial Intelligence*.
<https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence>

United Nations Educational, Scientific and Cultural Organization. (n.d.). *Readiness assessment methodology (RAM)*. UNESCO. <https://www.unesco.org/ethics-ai/en/ram>

United Nations General Assembly. (2024). *Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development (A/RES/78/265)*. United Nations Digital Library. <https://digitallibrary.un.org/record/4043244>

United Nations General Assembly. (2024). *Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development (A/RES/78/265)*. <https://documents.un.org/>

United Nations General Assembly. (2025). *Resolution 79/325: Independent International Scientific Panel on Artificial Intelligence and Global Dialogue on Artificial Intelligence Governance (A/RES/79/325)*. https://www.un.org/pga/wp-content/uploads/sites/110/2025/12/A_RES_79_325-EN.pdf

Working group on data governance at all levels. (2023). *UN Trade and Development (UNCTAD)*. <https://unctad.org/topic/commission-on-science-and-technology-for-development/working-group-on-data-governance>

World Summit on the Information Society. (2003). *WSIS declaration of principles. International Telecommunication Union*. <https://www.itu.int/net/wsis/docs/geneva/official/dop.html>

Xinhua News Agency. (2025). *China unveils Global AI Governance Action Plan*. <https://english.news.cn/>

Yinuo, & Yinuo. (2026). *Media Advisory | UN Forum to showcase practical, inclusive and scalable solutions to pressing global challenges*. United Nations Sustainable Development. <https://www.un.org/sustainabledevelopment/blog/2026/05/117750/>