January 2026
Florian Walta, President
Haganum Model United Nations XVI

Developing international frameworks for AI regulation

Group of 20 (G20)

Table of Contents

Introduction

Overview of the issue

Artificial intelligence (AI) has rapidly emerged as one of the most transformative technologies of the twenty-first century, reshaping economies, governance structures, and social interactions across the globe. From healthcare diagnostics and financial systems to military applications and public administration, AI systems are increasingly embedded in decision-making processes that affect billions of people. While these technologies offer significant opportunities for innovation and efficiency, their accelerated development has far outpaced the establishment of comprehensive international regulatory frameworks, creating a growing governance gap at the global level.

The absence of harmonised international regulations has resulted in fragmented national approaches to AI governance, often reflecting differing political priorities, economic capacities, and ethical standards. This lack of coordination increases the risk of regulatory arbitrage, in which companies or states exploit weaker oversight mechanisms, potentially leading to unsafe or unethical AI deployment. Moreover, AI systems trained on biased data or deployed without transparency can reinforce existing inequalities, violate privacy rights, and undermine trust in public institutions. These risks are particularly pronounced in cross-border contexts, where the impacts of AI technologies are not confined within national jurisdictions.

Impact on international security and global equity

The global nature of AI development also raises significant security and stability concerns. Autonomous weapons systems, mass surveillance technologies, and AI-driven cyber capabilities pose challenges to international peace and security if left unregulated. Without shared norms and accountability mechanisms, the misuse of AI could exacerbate geopolitical tensions, contribute to arms races, and erode established principles of international humanitarian law. Developing countries are especially vulnerable in this landscape, as they may lack both the regulatory capacity and technological infrastructure to safeguard their populations against harmful AI applications.

Need for international cooperation

Addressing these challenges requires the development of inclusive and adaptable international frameworks that balance innovation with ethical responsibility, human rights protection, and global equity. Effective international cooperation on AI regulation is essential to ensure that the benefits of artificial intelligence are distributed fairly, while minimizing its potential harms. By establishing shared standards, promoting transparency, and fostering capacity-building across states, the international community can work toward a future in which AI serves as a tool for sustainable development rather than a source of instability or inequality.

Definition of key terms

Algorithm

A set of rules or instructions designed to perform calculations, make decisions, or solve problems, often used by computer systems to process data and produce outputs.

Artificial intelligence (AI)

The capability of a machine or software system to perform tasks that normally require human intelligence, including learning, reasoning, pattern recognition, and decision-making.

Autonomous systems

Technological systems capable of operating and making decisions without direct human control once activated.

Bias (algorithmic bias)

Systematic and repeatable errors in computational systems that result in unfair or unintended outcomes, often arising from biased data or design processes.

Data privacy

The protection of personal or sensitive information from unauthorized access, misuse, or disclosure during data collection, processing, and storage.

Ethical AI

The development and use of artificial intelligence systems in accordance with moral principles such as fairness, accountability, transparency, and respect for human rights.

Governance frameworks

Structured sets of rules, principles, and procedures established to guide decision-making, oversight, and accountability within a specific domain.

Human oversight

The involvement of human judgement in monitoring, supervising, or intervening in the operation of automated or intelligent systems.

Machine learning

A subset of artificial intelligence that enables systems to learn from data and improve performance over time without being explicitly programmed for each task.

General Overview

Context and evolution of artificial intelligence

Artificial intelligence (AI) refers to a broad range of technologies that enable machines to perform tasks traditionally associated with human intelligence. Over the past two decades, advances in computing power, data availability, and machine learning techniques have accelerated AI development and deployment worldwide. AI-driven systems are now used in sectors such as healthcare, transportation, education, finance, law enforcement, and public administration. According to various estimates, AI-related technologies contribute trillions of dollars annually to the global economy, underscoring their growing economic significance alongside their societal impact.

The development of AI is largely driven by private-sector innovation, academic research, and government investment, often across borders. As a result, AI systems are frequently trained, deployed, and used in multiple jurisdictions simultaneously. This transnational nature complicates regulatory efforts, as existing legal frameworks are predominantly national and were not designed to address rapidly evolving, data-driven technologies.

Diverging national approaches to AI governance

States have adopted differing approaches to AI regulation, reflecting variations in political systems, economic priorities, and societal values. Some countries prioritize innovation-friendly environments, relying on voluntary guidelines and industry self-regulation to avoid constraining technological progress. Others emphasize precautionary regulation, introducing binding legal requirements related to data protection, algorithmic accountability, and risk assessment. These divergent approaches have led to a fragmented global regulatory landscape.

The lack of international alignment can create legal uncertainty for developers and users of AI systems operating across borders. It may also encourage regulatory arbitrage, where AI development or deployment shifts toward jurisdictions with fewer restrictions. At the same time, states with limited technological or regulatory capacity may struggle to implement effective oversight mechanisms, potentially widening existing global inequalities in technological governance.

Ethical, social, and human rights considerations

AI systems can raise a variety of ethical and social concerns, particularly when used in high-stakes decision-making. Issues such as algorithmic bias, lack of transparency, and limited accountability have been observed in areas including recruitment, credit scoring, predictive policing, and access to public services. These challenges often stem from biased training data, opaque system design, or insufficient human oversight.

From a human rights perspective, AI deployment may affect rights related to privacy, non-discrimination, freedom of expression, and due process. For example, large-scale data collection and AI-powered surveillance technologies can increase the risk of intrusive monitoring, while automated decision-making systems may limit individuals' ability to contest or understand outcomes that affect them. Addressing these concerns requires careful consideration of how existing international human rights norms apply to emerging technologies.

Security and military dimensions

AI also plays an increasing role in security and military contexts, including intelligence analysis, logistics, cyber operations, and weapons systems. Of particular concern is the development of autonomous weapons systems capable of selecting and engaging targets with limited or no human intervention. While proponents argue that such systems could reduce human error and improve operational efficiency, critics raise concerns about accountability, escalation risks, and compliance with international humanitarian law.

In addition, AI-driven cyber capabilities can enhance the scale and speed of cyberattacks, complicating attribution and response mechanisms. The absence of agreed international norms governing these applications raises concerns about arms races and unintended conflict escalation, highlighting the relevance of multilateral dialogue and confidence-building measures.

Economic implications and development gaps

The economic benefits of AI are unevenly distributed across and within states. Countries with advanced digital infrastructure, skilled workforces, and access to large datasets are better positioned to benefit from AI-driven growth. Conversely, developing countries may face barriers related to limited infrastructure, insufficient regulatory capacity, and dependency on foreign technologies.

At the same time, AI-driven automation has implications for labor markets, potentially displacing certain categories of jobs while creating demand for new skills. The pace and scale of these changes vary across sectors and regions, raising questions about social protection, education, and workforce adaptation. International cooperation may play a role in supporting capacity-building and knowledge-sharing to reduce these disparities.

Existing international initiatives and challenges

Various international and regional initiatives have sought to address aspects of AI governance through principles, guidelines, and non-binding frameworks. These efforts often emphasize values such as transparency, accountability, safety, and human-centered design. However, differences in scope, enforcement mechanisms, and participation limit their effectiveness as comprehensive global solutions.

A key challenge in developing international frameworks lies in balancing flexibility with legal certainty. Overly rigid regulations may hinder innovation and adaptability, while purely voluntary measures may lack accountability. Additionally, reaching consensus among states with differing interests and levels of technological development remains a significant obstacle.

Areas of convergence and ongoing debate

Despite these challenges, there is growing convergence around the idea that some level of international coordination on AI regulation is necessary. Areas of potential agreement include risk-based regulatory approaches, the importance of human oversight, and the need for transparency in high-risk applications. Ongoing debates center on questions such as the scope of regulation, the role of binding international law versus soft-law instruments, and the distribution of responsibilities among states, private actors, and international bodies.

The development of international frameworks for AI regulation therefore represents a complex and evolving issue, encompassing technological, legal, ethical, economic, and security dimensions. A comprehensive understanding of these facets is essential for informed policy discussion and effective multilateral cooperation.

Major parties involved

China

China is a leading global actor in the development and deployment of artificial intelligence, with significant state investment in AI research, infrastructure, and commercial applications. The government views AI as a strategic technology essential for economic growth, social governance, and national security. China has promoted national AI standards and governance guidelines that emphasize state oversight, security, and social stability. In international discussions, China generally supports the development of global AI norms but stresses the importance of state sovereignty, non-interference, and flexibility in implementation. Historically, China has favored regulatory approaches that allow governments substantial control over data flows and AI deployment, reflecting broader positions on cyberspace governance. These priorities shape its cautious stance toward binding international regulations that could limit domestic policy autonomy.

European Union

The European Union has positioned itself as a prominent advocate for comprehensive and rights-based AI regulation. Building on its earlier leadership in data protection and digital governance, the EU has pursued a risk-based approach to AI oversight that categorizes applications according to their potential impact on fundamental rights and safety. Its interest lies in ensuring that AI development aligns with principles such as transparency, accountability, and human dignity, while also maintaining competitiveness in the global digital economy. Historically, the EU has sought to export its regulatory standards internationally through market influence, a strategy sometimes referred to as the “Brussels effect.” In international forums, the EU generally supports stronger multilateral frameworks and binding rules, particularly for high-risk AI systems.

Private technology companies

Private technology companies play a central role in AI innovation, as they develop, deploy, and commercialize many of the most advanced AI systems in use today. Their interests often focus on maintaining innovation-friendly regulatory environments that allow flexibility, rapid iteration, and global market access. Many companies support international coordination to reduce regulatory fragmentation and compliance costs, but they frequently prefer non-binding guidelines or industry-led standards over strict legal obligations. Historically, self-regulation and voluntary ethical principles have been promoted as alternatives to formal international law. At the same time, growing public scrutiny and legal risks have led some companies to engage more actively in policy discussions and to support clearer rules for high-risk AI applications.

United Nations bodies

Various United Nations bodies are involved in AI-related discussions from legal, ethical, developmental, and security perspectives. These entities are generally interested in facilitating dialogue, promoting shared principles, and ensuring that AI development aligns with existing international norms, including human rights and international humanitarian law. The UN's role is shaped by its mandate to balance diverse state interests and levels of technological development. Historically, UN bodies have relied on expert groups, reports, and non-binding frameworks to address emerging technologies. In the context of AI regulation, they often emphasize inclusivity, capacity-building for developing countries, and the avoidance of technological divides, while recognizing the challenges of reaching consensus on binding global rules.

United States of America

The United States is a major center of AI research and innovation, driven largely by private-sector companies, academic institutions, and defense-related research. Its approach to AI governance has traditionally emphasized innovation, economic competitiveness, and national security, with a preference for flexible regulatory mechanisms. The United States has supported the development of ethical principles and sector-specific guidelines rather than comprehensive federal regulation, although this position has evolved over time. In international discussions, the United States generally supports cooperation on AI safety and standards but remains cautious about binding international frameworks that could constrain technological leadership or commercial interests. Historical concerns about strategic competition and intellectual property protection continue to influence its position.

Timeline of events

1956 June

The term artificial intelligence is coined at the Dartmouth Summer Research Project, marking the formal beginning of AI as a field of study.

1997 May 11th

An AI system defeats a reigning world chess champion in a match, demonstrating the potential of machine-based decision-making in complex tasks.

2012 September

Advances in deep learning significantly improve image recognition accuracy, accelerating global interest and investment in AI technologies.

2016 March

An AI program defeats a professional Go player, highlighting rapid progress in machine learning and strategic reasoning capabilities.

2018 May 25th

A comprehensive data protection regulation enters into force in multiple jurisdictions, influencing global discussions on data governance and AI oversight.

2019 May

International principles for trustworthy and human-centered AI are published, emphasizing ethics, transparency, and accountability.

2020 October

An international body adopts recommendations on the ethical use of artificial intelligence, providing a global reference framework for AI governance.

2021 September

A high-level international dialogue on lethal autonomous weapons systems continues, reflecting growing concern over military applications of AI.

2023 October

States begin negotiating comprehensive legal frameworks for high-risk AI systems, signaling a shift toward more binding regulatory approaches.

2024 March

An international resolution calls for enhanced global cooperation on artificial intelligence to ensure safe, secure, and trustworthy development and use.

Relevant UN treaties and events

Universal Declaration of Human Rights (UDHR)

Foundational UN document establishing universal human rights principles, which later inform discussions on privacy, non-discrimination, and human dignity in the context of emerging technologies, 10 December 1948.

International Covenant on Civil and Political Rights (ICCPR)

Legally binding treaty outlining civil and political rights, including privacy and due process, frequently referenced when assessing the impact of automated and AI-driven decision-making, 16 December 1966.

Convention on Certain Conventional Weapons (CCW)

Framework convention regulating specific weapons deemed to cause excessive harm, under which discussions on lethal autonomous weapons systems take place, 10 October 1980.

UN Secretary-General's Strategy on New Technologies

Strategic document outlining a UN-wide approach to emerging technologies, including artificial intelligence, with an emphasis on values-based and human-centered development, 20 September 2018.

UNESCO Recommendation on the Ethics of Artificial Intelligence

First global normative instrument adopted by a UN body specifically addressing AI ethics, providing guidance on governance, human rights, and international cooperation, 23 November 2021.

A/RES/77/21; Information and communications technologies for sustainable development

General Assembly resolution emphasizing the role of digital technologies, including AI, in achieving sustainable development while calling for international cooperation and capacity-building, 28 November 2022.

A/RES/78/69; Safe, secure and trustworthy artificial intelligence

General Assembly resolution encouraging enhanced international cooperation on artificial intelligence governance and the alignment of AI development with international law and human rights, 5 December 2023.

Previous attempts to solve the issue

European Union – risk-based regulatory approach

The European Union has pursued a comprehensive regulatory framework for artificial intelligence based on risk classification. This approach differentiates between minimal-risk, high-risk, and prohibited AI applications, with corresponding obligations for developers and users. The initiative builds on earlier EU experience in digital and data governance and represents one of the most detailed attempts to establish binding rules for AI systems within a multi-state entity.

United States – sector-specific and principles-based governance

The United States has largely addressed AI governance through sector-specific guidelines, executive actions, and voluntary principles rather than a single comprehensive law. Federal agencies have issued guidance on AI use in areas such as finance, healthcare, and defense, while broader ethical principles emphasize innovation, safety, and accountability. This approach reflects a preference for regulatory flexibility and adaptability to technological change.

China – national standards and state-led oversight

China has implemented national AI governance measures through administrative regulations, technical standards, and ethical guidelines. These efforts focus on security, data management, and alignment with national development goals. Oversight is primarily exercised through state institutions, and AI governance is integrated into broader digital and cybersecurity strategies.

OECD – international AI principles

The Organisation for Economic Co-operation and Development has developed non-binding principles for trustworthy artificial intelligence, endorsed by multiple countries. These principles address transparency, robustness, accountability, and human-centered values, and have influenced national AI strategies beyond OECD membership. The initiative represents an early attempt at international coordination without legal enforceability.

UNESCO – global ethical framework for AI

UNESCO has adopted a global recommendation on the ethics of artificial intelligence, providing guidance on governance, human rights, and societal well-being. While not legally binding, the framework aims to assist states in developing national policies and encourages international cooperation and capacity-building, particularly for developing countries.

Private sector – self-regulation and ethical guidelines

Major technology companies and industry associations have introduced internal AI ethics boards, codes of conduct, and impact assessment tools. These initiatives seek to address public concerns and guide responsible AI development within corporate structures. Such efforts vary widely in scope and enforcement and are often complementary to public regulatory initiatives.

Multilateral discussions on autonomous weapons systems

Within existing international forums, states have engaged in ongoing discussions regarding the regulation of autonomous weapons systems. These talks focus on definitions, human control, and the applicability of existing international law, reflecting attempts to address AI-related security concerns through dialogue rather than new binding agreements.

Possible solutions

Establishing a non-binding global framework of AI principles
States could agree on a set of shared international principles guiding the development and use of artificial intelligence. Such a framework would focus on common values such as safety, transparency, human oversight, and respect for human rights, while allowing states flexibility in domestic implementation. This approach may encourage broad participation and serve as a foundation for future cooperation, though its effectiveness would depend on voluntary compliance and peer pressure mechanisms.

Developing binding international regulations for high-risk AI systems
The international community could negotiate legally binding rules targeting high-risk applications of AI, such as systems used in critical infrastructure, military contexts, or large-scale surveillance. This solution would aim to create minimum global standards and accountability mechanisms while leaving low-risk AI largely unregulated. Debate may arise over scope, enforcement, and the balance between innovation and precaution.

Creating an international AI oversight or coordination body
A dedicated international body could be established to facilitate information-sharing, monitor AI developments, and provide technical guidance to states. Such an institution could support harmonisation of standards, assist with capacity-building, and serve as a platform for dialogue on emerging risks. Questions may emerge regarding mandate, authority, funding, and overlap with existing international institutions.

Promoting capacity-building and technology transfer for developing countries
International cooperation could prioritize assistance for developing states in building regulatory, technical, and institutional capacity related to AI governance. This may include training, shared research initiatives, and access to best practices. Supporters argue this could reduce global inequalities, while critics may raise concerns about resource allocation and implementation feasibility.

Adopting a risk-based and sector-specific regulatory approach
Rather than a single comprehensive framework, states could coordinate sector-specific standards for AI use in areas such as healthcare, finance, transportation, and defense. This approach allows tailored regulation that reflects the unique risks and requirements of each sector. However, it may also increase fragmentation and complicate international interoperability.

Strengthening transparency and accountability mechanisms
States could require AI developers and deployers to conduct impact assessments, maintain documentation, and enable independent audits for certain AI systems. At the international level, shared guidelines for transparency could enhance trust and facilitate cross-border cooperation. Debate may focus on confidentiality, intellectual property, and compliance costs