ie UNIVERSITY
CENTER FOR THE GOVERNANCE OF CHANGE

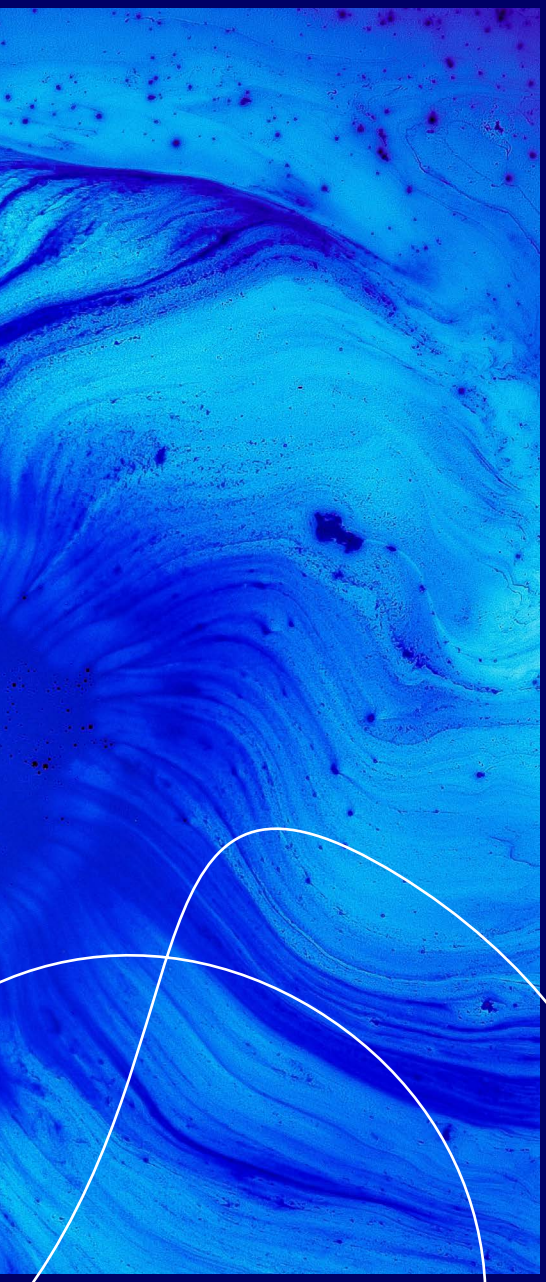
AI4 DEMOCRACY

JULY 2024

# THE **G20** AND GLOBAL AI GOVERNANCE

By Roxana Radu

# ABSTRACT



In the words of the UN Secretary-General on 12 June 2024[i],

**artificial intelligence (AI) is the biggest revolution since the invention of internet – perhaps since the second industrial revolution.**

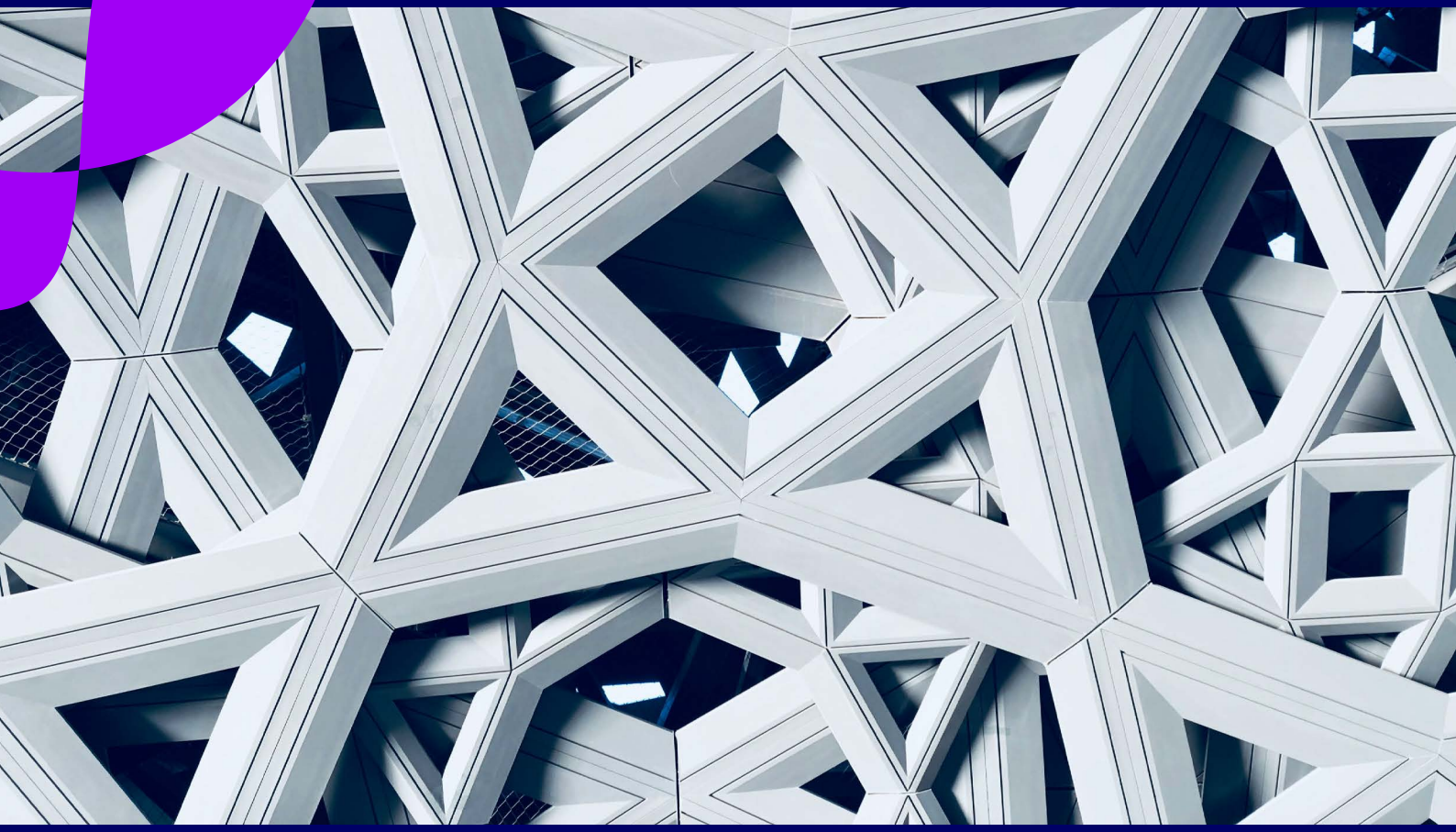
There is emerging consensus that its power is transformative, if harnessed "for good", though what the latter means is still open to debate. As a general-purpose technology, AI changes human interactions and permeates the fabric of society, heightening uncertainty. While long-termist scholars have highlighted the existential risks to humanity, many more have focused on the intended and unintended consequences and harms resulting from the deployment of AI in the short-term.

Numerous studies and policy reports have categorized these risks and opportunities, analysing various legislative measures, yet a comprehensive evaluation of emerging pathways toward global AI governance is currently missing. This white paper aims to fill this important gap by providing an evidence-informed analysis of nascent governance mechanisms, putting them in conversation with ongoing G20 efforts to explore paths forward.

# TABLE OF CONTENTS

# ACKNOWLEDGEMENTS

The author gratefully acknowledges the Center for the Governance of Change at the IE University for organising and hosting a dedicated meeting on 28 May 2024 in Madrid.



# ARTIFICIAL INTELLIGENCE IN A GLOBAL PERSPECTIVE: **2024**

Despite repeated calls at the highest level for a collective approach to governing Artificial Intelligence (AI), a fragmented approach has gained momentum. Evidence from both academic experts and international organisations[ii] points to this reality: AI governance efforts are disjoined and highly politicised. The emerging AI ecosystem has not (yet) matched the aspiration towards an inclusive, transparent and equitable global system. The control of AI infrastructure and critical resources has heightened, rather than reduced, geopolitical tensions. Major divides such as public vs private ownership, North vs global South and high capability vs low capability have structured the conversation thus far[iii]. In the absence of an agreed definition for AI, different governance models have emerged, providing a wide spectrum of both soft and hard law instruments, implemented unilaterally or in conjunction with various stakeholders. The emphasis on self-regulation and ethical principles, ever-present prior to 2020, has given way to serious discussions about regulatory approaches centred on risk mitigation and harm prevention.

THE EMERGING AI ECOSYSTEM HAS NOT (YET) MATCHED THE ASPIRATION TOWARDS AN INCLUSIVE, TRANSPARENT AND EQUITABLE GLOBAL SYSTEM.

The latest advancements in deep learning and neural networks, popularly referred to as 'artificial intelligence', have accelerated this trend. Post-2020, AI tools – such as text generation and image recognition – have become widely integrated into everyday use applications, be they anti-virus software or virtual assistants. However, the rise of new generative AI models – such as ChatGPT or Gemini – has changed the digital landscape. This new wave of AI represents a shift towards a general-purpose technology, integrating and producing different types of content, including text, images, video, audio and code. It therefore brings about a growing number of decisions embedded in software code and delegated to self-learning algorithms. Additionally, these AI systems are incredibly versatile, capable of being applied to a wide range of contexts and serving various purposes, making them powerful and unpredictable in their applications.

—

BY 2023, THE NUMBER OF PROJECTS USING AI TRIPLED

Existing AI systems can already proliferate rapidly and learn autonomously from the billions of data points made available to them. They have thus changed the conversation: they show potential for increased productivity or medical advances, but they have already started to displace workers, increase surveillance and further inequality and discrimination. In many cases, even the developers who design these AI systems may not fully grasp how they function or generate their results. To address this new set of challenges, many countries have gradually built on their national strategies[iv] to create parliamentary bills on AI. Such developments echo a push towards legislation, materialising in multi-year negotiation processes at the level of the European Union (EU) Artificial Intelligence Act and the Council of Europe (CoE) Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law.

While AI is still in its infancy, decision-makers at all levels are confronted with hard choices regarding its governance and its daily use. The ambition to deploy it rapidly to foster sustainable development, social justice and wellbeing faces the reality of managing associated risks and minimising harms. Within and outside of the UN, the world's largest international organisation, its member states and bodies have shown enthusiasm and optimism for deploying digital technologies to achieve the Sustainable Development Goals (SDGs). In 2019, more than 35 UN main bodies, specialized and related agencies experimented with the deployment of AI in their activities, ranging from predictive analysis deployed in their operations to designing standards, policies and normative frameworks for future AI developments[v]. By 2023, the number of projects using AI tripled[vi]. Around the world, AI adoption in the public and private sectors has gained considerable traction, whether as a support tool in public administration or a competitive advantage in the market.

The market itself has a key role to play in the distribution of AI benefits and power. The core dependency in scaling-up AI has been access to computing power, alongside data and skilled labour. The computational infrastructure, both hardware and software (from semiconductors to data management applications) is currently highly concentrated at key points in

THE G20 CAN PLAY A SIGNIFICANT ROLE IN STEERING THE GLOBAL GOVERNANCE OF AI

the supply chain[vii] (Data and Society 2023). Major tech companies that dominate the cloud services market possess a substantial advantage in developing and controlling the infrastructure essential for AI projects. Their dominance allows them to set the standards and pace of innovation in AI, providing them with unparalleled influence over who can access top-tier computational resources. This control not only reinforces their market position but also shapes the direction of AI research and development. Additionally, the availability of large datasets, crucial for training AI models, is often limited to entities with extensive data collection capabilities. Variations in data privacy regulations further complicate matters, affecting who can collect and utilize data effectively.

Against this background, the digital divides persist and new gaps are created, as countries and regions lacking adequate access to computational infrastructure risk lagging behind in the development of AI. Unequal access to computational resources limits these regions' capacity to participate fully in the AI-driven digital economy, to develop critical expertise and to harness indigenous innovation efforts, perpetuating disparities on a global scale.

Positioned as one of the most significant intergovernmental forums spanning five continents, the G20 can play a significant role in steering the global governance of AI. Consisting of 19 sovereign states plus the European Union and recently admitted African Union, the G20 represents two-thirds of the world's population and accounts for 85% of global GDP. Brazil currently holds the G20 presidency, leading efforts to coordinate technical meetings, ministerial conferences and the upcoming summit of heads of state scheduled for November 18-19, 2024 in Rio de Janeiro. High on Brazil's agenda is the focus on Artificial Intelligence for Sustainable Development and Reduction of Inequality, aiming to address the societal and political impacts of increasingly potent digital technologies.

# THE INTERNATIONAL AI GOVERNANCE ECOSYSTEM

EFFORTS TO GOVERN AI CAN BE COMPARED TO MANAGING TRAFFIC AT A BUSY CROSSROADS

Current international discussions underscore the necessity for a unified framework to address the global implications of AI technologies. These technologies present worldwide challenges that demand a concerted response. However, the ongoing debate centres around the form this harmonised approach should take and the extent to which it serves the stakeholders' interests. From specific calls to ban AI development to proposals for building AI as public infrastructure, no clear-cut path has emerged on advancing AI for the benefit of all. There remains divergence of opinion on whether a unified global approach risks becoming too uniform and failing to accommodate the diverse needs and contexts across different regions and sectors, maintaining power imbalances. Many of these critical deliberations persist in silos, mirroring distinct perspectives from technical, legal, political and cultural traditions.

Moreover, the lack of clear international coordination frameworks exacerbates the challenge of managing AI as a global, general-purpose technology, used in both military applications and humanitarian support for children in need[viii]. This situation leads to a critical implementation challenge: should we rely on a weak regime with minimal enforcement power or strive for a single, robust institution to oversee AI governance? The debate reflects a host of answers given so far, on a broader spectrum of governance mechanisms, from soft law (relying on voluntary compliance and flexibility) to hard law (which entails binding regulations and strict enforcement).

Efforts to govern AI can be compared to managing traffic at a busy crossroads, where different generations of vehicles are meeting. Currently, some discussions focus on individual safety measures, similar to ensuring seat belts are in place within the cars and drivers of two-wheelers wear protection helmets. Addressing the potential harms of AI systems during deployment is crucial for ensuring their safe and responsible use. This approach can be compared to fixing a car while driving it,

emphasizing the need to mitigate risks in real-time and prioritize the safety and well-being of users. These concerns revolve around establishing robust safeguards and regulatory frameworks to prevent AI systems from causing unintended harm or being misused.

Others are concerned with maintaining control over AI systems, questioning if these technologies are becoming akin to driverless cars. The key issue here is how to keep humans involved in the decision-making processes of AI systems. This involves designing AI systems that allow for human oversight and intervention, ensuring that humans retain ultimate control over critical decisions. Additionally, there is a growing need to envision meaningful roles for humans in an era where job displacement due to automation and AI is a significant concern. This involves not only creating new job opportunities but also ensuring that humans can work alongside AI in a complementary manner, enhancing productivity and job satisfaction.

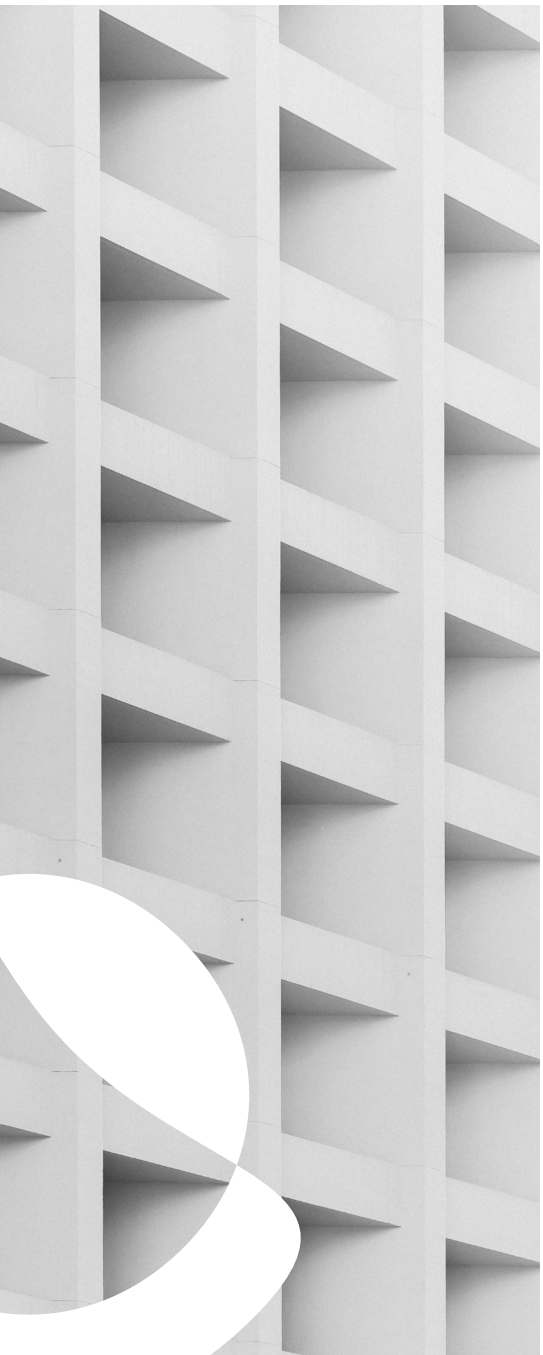
A third category of concerns revolves around existential risks posed by AI. These concerns delve deeper into the long-term implications of AI development, questioning whether AI systems could threaten human autonomy and control. Some fear we may be approaching a dystopian future where AI systems and robots dominate humanity, leading to a loss of human agency and potential apocalyptic scenarios. This perspective often captures the public imagination and dominates much of the media portrayal of AI. However, while these sensationalized views attract attention, they often overshadow more immediate and practical issues that concern the design and deployment of safe AI systems.

Despite the varied focus of these concerns, there is a consensus that both short-term and long-term challenges associated with AI require effective governance. Currently, there is no unified global governance system for AI. Instead, what is emerging is a diverse governance ecosystem characterized by varying national preferences and uneven institutional capacities. Different countries are adopting their own approaches to AI regulation, leading to a fragmented landscape. Some nations prioritize innovation and economic growth, while others emphasize ethical considerations and human rights.

This divergence highlights the need for international cooperation and harmonization of AI governance frameworks to ensure that AI development benefits all of humanity equitably.

If global governance is understood as 'a set of authoritative rules aimed at defining, constraining and shaping actor expectations in a purposive order, generally implemented through a set of mechanisms recognised as legitimate by relevant actors'[ix], it is worth considering efforts at the multilateral level alongside influential initiatives from non-state actors. Numerous initiatives emerging post-2019 – whether in the forms of codes of conduct, principles, guidelines – formalise commitments to raise the bar for protections afforded in an AI-driven future[x]. Many of them are public, though some may remain behind closed doors. A recent example is the summit between the Chinese President Xi Jinping and U.S. President Joe Biden in San Francisco at the end of 2023, which paved the way for an intergovernmental dialogue on AI started in Geneva in May 2024, on the side-lines of the AI for Good Summit[xi].

VARIOUS INTERNATIONAL ORGANIZATIONS AND BODIES HAVE STARTED WORKING ON THE BUILDING BLOCKS FOR AN INTERNATIONAL GOVERNANCE SYSTEM.

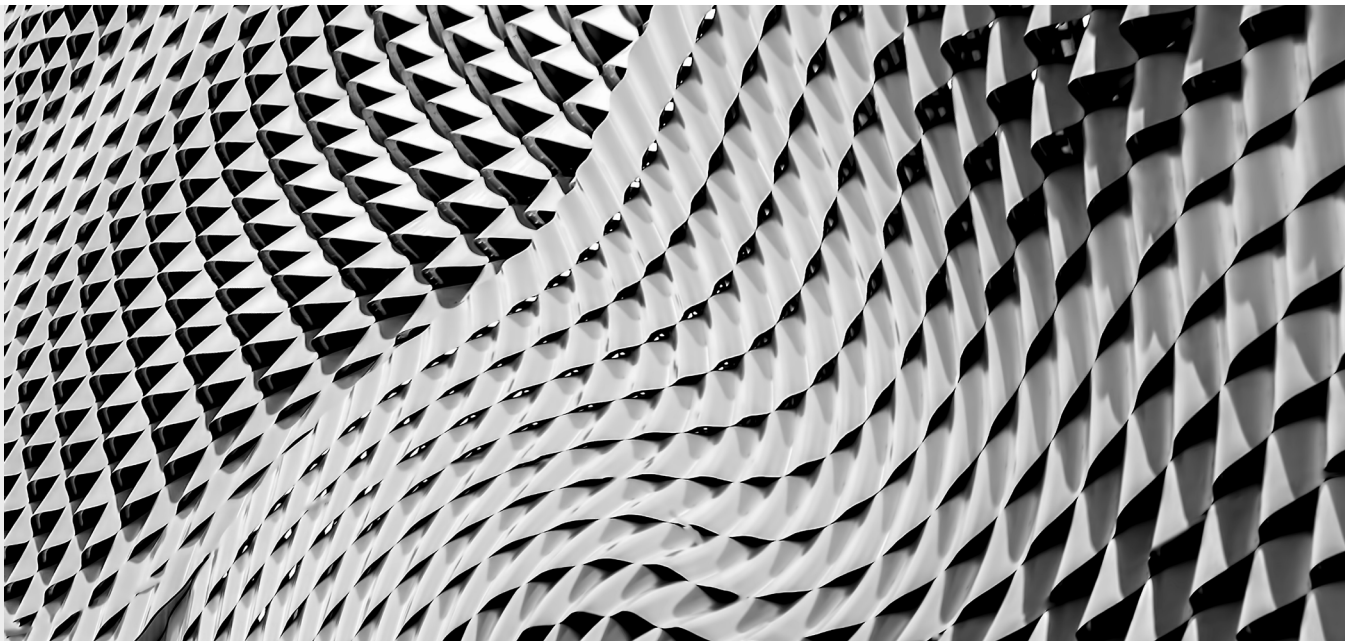


Various international organizations and bodies have started working on the building blocks for an international governance system. The AI for Good Summit was one of the early efforts – initiated by the International Telecommunication Union in 2017 – aiming to bring together policymakers, technologists, academia and civil society to Geneva to explore and promote the use of artificial intelligence for societal benefit. It continues as an annual summit focused on leveraging AI technologies to address global challenges outlined in the United Nations Sustainable Development Goals (SDGs), including healthcare, education, climate change and inequality. With a different focus, the Global Partnership on AI (GPAI) has focused on facilitating collaboration among governments, international organizations and private sector entities. But the AI governance landscape has significantly diversified since, with global platforms and events occurring on a monthly basis all over the world.

# DECODING GLOBAL AI GOVERNANCE INITIATIVES: A FRAMEWORK FOR ANALYSIS

Given the wide variety of proposals for governing AI, it is necessary to dissect actions at the regional and international level, where states and non-state actors work together to achieve a more coherent AI ecosystem. There are wide-ranging and highly diverse initiatives, ranging from global agencies proposals to narrow structures and regulations at the local level (e.g. New York audit law), which complement hybrid governance approaches prioritised at the national level (Radu, 2021) and standard-setting processes initiated by technical bodies[xii]. Some of the normative foundations – even when agreed by a majority – suffer from limited implementation thus far. Repeatedly, AI has been discussed as part of broader concerns in relation to emerging technologies, although AI-specific governance instruments are starting to proliferate.

To unpack the complexity of the emergent AI landscape, a categorisation of governance mechanisms is much needed. A tripartite framework based on my earlier work on internet governance[xiii] is applied here. This framework distinguishes among three dominant mechanisms, considered on a soft-hard law spectrum. Legally-binding instruments and soft law instrument sit at the different end of the spectrum, with institutional solidification as the middle ground.



**The three mechanisms can be divided as follows:**

## 1. MODELLING

Modelling involves non-binding commitments that are formally announced to indicate a stakeholder's position or to influence and prompt actions by other actors, such as setting industry standards. There are two main types of modelling:

1

**Discursive Modelling**: Relies on statements and declarations.

2

**Operative Guidance Tools**: Includes practical recommendations, guidelines and model documents.

Modelling instruments differ from legal enshrinement mechanisms by offering more flexibility and resulting from shorter negotiation processes. These mechanisms are open to all actors, including technical bodies, academia, civil society groups and corporate players.

## 2. INSTITUTIONAL SOLIDIFICATION

Positioned between legal enshrinement and modelling, institutional solidification involves efforts to make a procedure, working group or concern more permanent. This mechanism often involves creating a structured institutional design, such as a dedicated secretariat or a person on payroll, leaving a trace in global governance. Institutional solidification can be categorized by intentionality and scope of actions:

1

**Specialized Bodies**: Expert work coordinated in a structured manner (e.g., UN Advisory Body on AI).

2

**Strategic Frameworks, Plans of Action and Global Agendas**: Set objectives for collective work, often with attached funding.

3

**Monitoring and Benchmarking Tools**: Require iterative processes (e.g., annual rankings, global database updates) and long-term financial commitment.

## 3. LEGAL ENSHRINEMENT

Legal enshrinement refers to binding instruments that are initiated and signed by states or issued by public authorities such as courts or supra-national bodies. These instruments, which include treaties, conventions and related agreements like protocols, have a particular status under international law and are exclusively open to state actors. Conversely, court judgments, directives and binding policies and legislation can target private actors while maintaining the requirement for compulsory action.

The analysis of 28 global initiatives presents us with the following picture:

| Governance Mechanisms | Instruments | Instances |
|---|---|---|
| **Modelling** | Discursive actions | 2019/2024 OECD AI Principles<br>2021 African Commission resolution 473<br>2023 G20 New Delhi Leaders' Declaration<br>2023 Bletchley Declaration<br>2024 UN General Assembly resolution on AI |
| | Operative guidance | 2021 UNESCO Recommendation on the Ethics of Artificial Intelligence<br>2023 G7 Hiroshima Process on AI Guiding Principles<br>2023 Voluntary White House AI Commitments<br>2024 ASEAN Guide on AI Governance and Ethics |
| **Institutional solidification** | Specialised bodies | 2019 AU working group on AI<br>2023 UN AI Advisory Body<br>2023 BRICS "AI study" group<br>2024 EU AI Office |
| | Strategic frameworks/ action plans | 2021 AI for Africa Blueprint<br>2023 Global AI Governance Initiative |
| | Monitoring and benchmarking tools | OECD AI Policy Observatory<br>Stanford AI Index<br>Global AI Index<br>AI Global Surveillance Index<br>Artificial Intelligence and Democratic Values Index<br>AI Incident database<br>AI readiness Index<br>Global Index on Responsible AI<br>African Observatory on Responsible AI<br>Latin American Artificial Intelligence Index (2023)<br>AI Governance International Evaluation Index |
| **Legal enshrinement** | Treaties, conventions, binding agreements | 2024 CoE Framework Convention on AI and human rights, democracy and the rule of law |
| | Legislation, directives with global/regional effects | 2024 EU AI Act |

Over time, the initial emphasis on soft law instruments and AI principles has paved the way for more tangible AI governance proposals geared towards institutional action and hard law. A case for AI diplomacy has recently been made[xiv], as this policy starts to consolidate on the international agenda. While a basic vocabulary has been developed in the process of developing shared values and principles, there is a long way to go for agreeing the forms that AI governance and regulation might take and what degree of agility and adaptiveness is needed to respond to new technological developments.

## MODELLING

The 36-member Organization for Economic Co-operation and Development (OECD) adopted a set of AI ethics principles aimed at guiding the development and deployment of AI technologies in 2019 and revised them in 2023. These principles emphasise values such as transparency, accountability and human rights, serving as a foundation for ethical AI practices. The G20 – comprising the world's major economies – committed to these principles that same year, reflecting a broad consensus on the importance of ethical AI governance. By the same token, in November 2021, all 193 UN member states adopted the UNESCO Recommendation on the Ethics of Artificial Intelligence, committing to ethical principles in AI development. However, until now, less than a quarter of its signatories have implemented the proposed policy tools. Finally, in 2024, the UN General Assembly adopted a resolution on AI – "Seizing the opportunities of safe, secure and trustworthy AI systems for sustainable development" – underscoring the global community's dedication to advancing AI governance and ensuring that AI technologies contribute positively to society. These collective efforts highlight the international community's intention to proactively manage AI's transformative impact through collaborative and ethical governance frameworks, but also pinpoint the limits of non-binding agreements, whether in global configurations or in narrower membership organisations such as the GPAI, where no agreement has been reached on the governance initiatives to pursue.

More progress seems to have been made at the regional level, as well as in the club format. Beyond the EU, which has issued the first binding rules for its 27 member states, in the African Union and in the ASEAN progress has been made through modelling mechanisms. The African Commission's Resolution 473, dating back to February 2021, called for regulation grounded in the needs of Africans and serves as a reference for national legislation. The latter function is also reflected in the Asian approach, as underscored by the 2024 ASEAN Guide on AI Governance and Ethics.

In 2023, the G7 initiated the Hiroshima AI Process to enhance international cooperation in AI governance, reflecting the growing recognition of AI's global impact and the need for collaborative governance frameworks. The process has also defined an International Code of Conduct for Developers of Advanced AI System and the OECD has started designing a mechanism to monitor its implementation for those who choose to adopt it.

Globally, the G20 Leaders' Declaration from 2023 was a decisive moment to draw attention to the need for a harmonised, yet differentiated approach, centred on "Harnessing AI responsible for good and for all", in particular with respect to promoting responsible AI for achieving the SDGs and in support of the digital economy. Reaffirming the commitment to the G20 AI Principles (2019), the Declaration also stressed the role of the Data for Development alignment.

—

THE G20 LEADERS' DECLARATION FROM 2023 WAS A DECISIVE MOMENT TO DRAW ATTENTION TO THE NEED FOR A HARMONISED, YET DIFFERENTIATED APPROACH

On the high-risk AI front, the United Kingdom established and AI Safety Summit (hosted at Bletchley Park in November 2023) to advance global knowledge on advanced AI, focusing on safety and ethical considerations. Its second edition took place in Seoul in May 2024 and saw the launch of the first international network of AI Safety Institutes and a new commitment from 27 governments to deepen work on establishing thresholds for severe AI risks and AI safety testing and evaluation guidelines. In parallel, 16 AI companies from the US, China, Middle East and Europe signed the "Frontier AI Safety Commitments" promising to publish their safety framework(s) ahead of the AI Action Summit in France in 2025.

Despite the development of authoritative AI principles, the rather abundant soft law efforts remain fragmented across various domains and groupings, compartmentalising the global conversation. This siloed approach has already hampered the potential for a more cohesive strategy for AI governance and continues to have reverberations across the globe.

## INSTITUTIONAL SOLIDIFICATION

In 2023, many analogies with existing governance structures were discussed by both scholars and policymakers, including the International Civil Aviation Organization (ICAO), the European Organization for Nuclear Research (CERN), the International Atomic Energy Agency (IAEA), the Intergovernmental Panel on Climate Change (IPCC), the Financial Action Task Force (FATF) and the Financial Stability Board (FSB). While the best institutional form for a global body continues to be debated, coordination efforts have advanced on multiple fronts. The Trade and Technology Council was established to coordinate EU and US activities in trade and technology, including AI, promoting transatlantic cooperation on emerging technologies. With discussions at an early stage, the BRICS nations agreed to form an 'AI study group,' reinforcing their collective interest in shaping the future of AI. An analysis of declarations from leaders' summits and ministerial meetings indicates that the BRICS nations predominantly view AI as a key driver for economic growth, development, technological progress and the creation of inclusive societies. Statements from these gatherings frequently emphasize the importance of collaboration and progress in the field of AI, highlighting it as a crucial component of their strategic agendas[xv].

WHILE THE BEST INSTITUTIONAL FORM FOR A GLOBAL BODY CONTINUES TO BE DEBATED, COORDINATION EFFORTS HAVE ADVANCED ON MULTIPLE FRONTS.

At the UN level, the Secretary-General appointed a High-Level Advisory Body on AI to advance recommendations for international AI governance, ahead of the Summit of the Future scheduled for 22-23 September 2024. The group released its interim report at the end of last year and consulted extensively ahead of the final report, due to be published in the summer of 2024.

In the AI landscape, purpose-built initiatives are more and more common, but so are the critiques. AccessNow's decision to

withdraw from the Global Partnership on AI (GPAI) is one such example. AccessNow, a global NGO, cited the excessive influence of private sector interests within GPAI as a key reason for their withdrawal, highlighting how such dominance can undermine the credibility and effectiveness of these platforms[xvi].
Relatedly, the benchmarking exercises constitute the large majority of initiatives in this space. Out of a total of 28 initiatives, 11 perform the observatory and index roles, suggesting the monitoring of developments on both the technological and the policy side are in high demand, in particular for comparing the regional dynamics. Since many of them started in late 2023 and early 2024, it is difficult to estimate how sustainable they are beyond the initial funding cycle.

## LEGAL ENSHRINEMENT

The Council of Europe has actively pursued the creation of a global regulatory framework for artificial intelligence (AI) since 2019 through its Ad Hoc Committee on Artificial Intelligence (CAHAI). This effort involved extensive assessments to explore the feasibility of establishing a legal framework that covers the development, design and application of AI technologies. In June 2022, the Committee of Ministers of the Council of Europe tasked the new Committee on Artificial Intelligence (CAI) with negotiating an international agreement for AI. This mandate emphasized the importance of upholding the Council's existing norms while fostering innovation in AI technologies. The goal was to ensure that any regulatory framework not only supports technological advancement but also protects fundamental rights and maintains democratic principles.

The Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law was adopted in Strasbourg on 17 May 2024. This landmark convention imposes obligations on all future parties to mitigate risks associated with AI activities conducted by both public and private actors throughout the AI lifecycle. It emphasizes the distinct roles and responsibilities of various stakeholders and allows parties flexibility in meeting these obligations within their own legal and institutional frameworks.

THE EUROPEAN UNION HAD ITS FIRST DISCUSSIONS ABOUT AI IN 2015 AS PART OF A EUROPEAN PARLIAMENT EXPERT GROUP EXAMINING CIVIL LAW RULES ON ROBOTICS.

A key feature of the convention is its periodic reporting mechanism, requiring signatories to report on the measures they have implemented. This mechanism aims to enhance state accountability and facilitate ongoing adaptation to the evolving landscape of AI technologies. Additionally, the convention includes a follow-up mechanism designed to foster cooperation with states that have not yet ratified the treaty, potentially expanding its global influence and effectiveness.

Concurrently, the European Union had its first discussions about AI in 2015 as part of a European Parliament expert group examining civil law rules on robotics. Alongside many other efforts, the EU has been developing its AI legislative package since 2021, building on two pillars. The first was a regulatory one, culminating in the adoption of the EU AI Act in May 2024, with provisions tailored to oversee both general-purpose AI systems and high-risk AI applications. Alongside the risk-based regulation and product safety approach, the EU has also developed a long-term vision for research innovation and investment in AI, to increase computing power and access for small and medium-sized enterprises. In the European Union context, these enablers and binding rules build on existing regulations around data protection, privacy and consumer protection, regulating AI in a more comprehensive manner within the EU internal market, while safeguarding fundamental human rights and democratic values.

With these important initiatives across the CoE and the EU, the European region has become the first to adopt legally-binding rules that span an important number of countries.

# THE ROLE OF THE **G20**

AS A FORUM UNITING THE WORLD'S MAJOR ECONOMIES, THE G20 HAS THE INFLUENCE AND SCOPE NEEDED TO BRING COHERENCE TO THE HIGHLY FRAGMENTED AND IMBALANCED AI ECOSYSTEM.

At a pivotal moment in AI's evolution, the G20 is uniquely positioned to advance the global agenda for this new technology. The interconnected challenges confronting G20 nations, whether technological or developmental, underscore the necessity for effective and holistic policy coordination. As a forum uniting the world's major economies, the G20 has the influence and scope needed to bring coherence to the highly fragmented and imbalanced AI ecosystem. The Brazilian presidency of the G20 is crucial for amplifying the voice of the Global South, especially following India's tenure (2023) and preceding South Africa's leadership (2025).

During a dedicated IE meeting in Madrid on 28 May 2024, participants from the diplomatic community and major stakeholder groups recognised AI as an unprecedented challenge and a watershed moment, emphasising its multifaceted impact on society. They stressed the necessity of a holistic approach to AI, considering not only technological advancements but also the ethical, social and human implications. This holistic perspective is essential to address the wide-ranging consequences of AI, both positive and negative. Building on its early adoption of AI principles in 2019 and the momentous 2023 Delhi Declaration, the G20 has established a strong foundation for steering the governance of AI. The G20 is committed to a human-centric approach, empowering sustainable digital transformation and fostering inclusive growth. However, developing countries face significant challenges, such as the risk of furthering power inequalities in the absence of a transparent and inclusive governance ecosystem. Current governance efforts do not reflect the needs of the Global South, and there is a pressing need to integrate their voices into the decision-making process to ensure equitable development. The G20 Digital Economy Working Group, with its diverse approaches to economic development, can thus serve as a platform for dialogue and consensus-building. Brazil's presidency supports equitable access to AI, emphasizing sustainability and inclusivity, which are crucial for the Global South.

Key themes from the discussion in Madrid included the G20's crucial role in facilitating international cooperation and the importance of public-private collaboration. Participants highlighted the need to blend regulation with innovation to harness AI's potential for sustainable and inclusive development, in particular as computing power remains a scarce resource. The integration of AI and cybersecurity was also deemed vital, requiring secure and interoperable tools to ensure safe digital environments. Digital inclusion remains the starting point for connectivity and safety, as highlighted by some of the participants.

ETHICAL AND RESPONSIBLE AI USE, EMPHASIZING TRANSPARENCY, FAIRNESS AND ACCOUNTABILITY, CAN HELP ADDRESS THE SHORTCOMING OF STANDALONE PRIVATE SECTOR SELF-GOVERNANCE.

Ethical and responsible AI use, emphasizing transparency, fairness and accountability, can help address the shortcoming of standalone private sector self-governance, as recently seen with generative AI. International cooperation is essential in order to harmonise regulatory frameworks, and both multilateral and multistakeholder approaches are necessary for effective governance. The G20, as a consensus-based organisation, can move the global discussion forward, supporting existing efforts at the UN level – such as the work of the UN Advisory Body on AI – and bringing coherence to the fragmented AI governance landscape. Advancing the Delhi Declaration's principles, the group can address key deficits in normative convergence, while working to identify commonalities and strengthening a global good approach.

# RECOMMENDATIONS

The numerous calls for global AI governance have led to very different responses, ranging from modelling the behaviour of others to developing full-fledged regulatory frameworks. Based on the wide variety of interests it represents and its convening power, the G20 can help streamline the various ongoing processes and minimize unnecessary duplications. Looking ahead, under the Brazilian presidency of the G20 and beyond, there are three actionable directions that can have a substantial impact on the evolving AI ecosystem:

**I ) Strengthening and operationalizing existing commitments, in particular:**

- Leading by example in enhancing transparency in the AI landscape, through open, clear and accessible policies and practices and regular updates on the G20 actions for different stakeholder communities.
- Minimizing the environmental footprint of AI systems by adopting sustainable approaches throughout the AI life cycle, from prioritizing energy-efficient AI research and development to implementing standards for green data centers and electronic waste recycling.

**II ) Building a common horizon for AI governance, defining how a responsible and human-centred approach can help overcome current tensions in establishing a global regime for AI, by:**

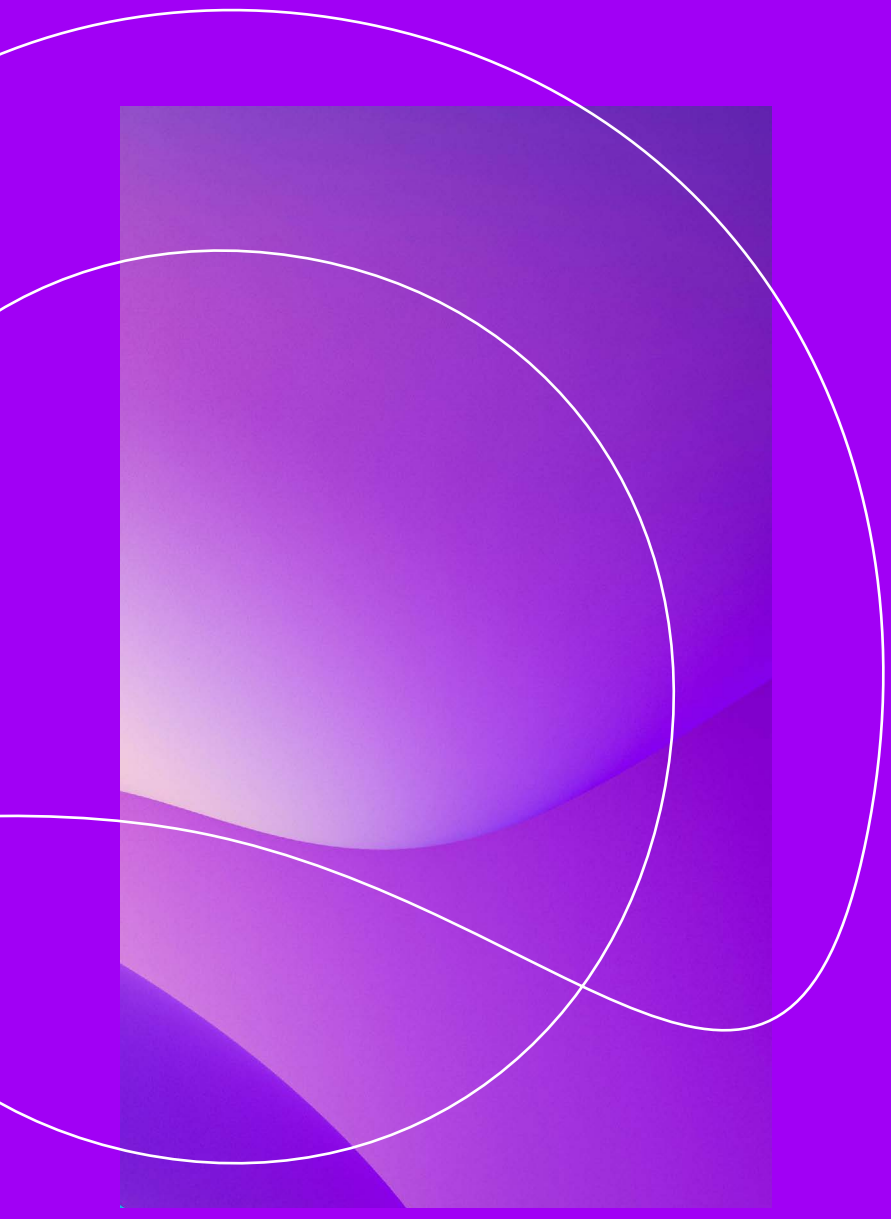
- Developing a forward-looking research agenda, to better articulate how the AI ecosystem evolves through interactions with current and emerging technologies.
- Creating a global coordination plan, to ensure alignment with UN processes for the cohesiveness and unity in AI governance efforts

## III ) Charting new paths for global AI discussions by:

**1.** Proposing a remedy framework for AI harms, that members of the G20 can take forward in their national legislation

**2.** Giving a voice to future generations likely to be affected by the evolution of AI and committing explicitly to upholding children rights in AI mainstreaming.

**3.** Prioritising work at the level of standards and certification, aiming towards a harmonised framework and a public forum for open discussions



# ENDNOTES

i ITU (2024). UN Secretary General remarks to the ITU Council (12 June). Available at: https://www.itu.int/hub/2024/06/un-secretary-general-remarks-to-itu-council/

ii Butcher, J., & Beridze, I. (2019). What is the State of Artificial Intelligence Governance Globally? RUSI Journal, 164(5–6), 88–96; Veale, M. M. (2023). AI and Global Governance: Modalities, Rationales, Tensions. Annual Review of Law and Social Science 19; Ulnicane, I. and Erkkilä, T. (2023), Politics and policy of Artificial Intelligence. Rev Policy Res, 40: 612-625. Erman, E., & Furendal, M. (2024). Artificial Intelligence and the Political Legitimacy of Global Governance. *Political Studies*, 72(2), 421 -441.

iii Radu, R. (2024). The variable geometry of AI governance. Geneva Policy Outlook. Available at: https://www.genevapolicyoutlook.ch/the-variable-geometry-of-ai-governance/

iv Radu, R. (2021). Steering the governance of artificial intelligence: national strategies in perspective, *Policy and Society*, Volume 40, Issue 2, June 2021, Pages 178–193, https://doi.org/10.1080/14494035.2021.1929728

v ITU (2019). UN Activities on Artificial Intelligence. Available at: https://www.itu.int/hub/publication/s-gen-unact-2023/

vi idem

vii Vipra, J. and Myers West, S. (2023). Computational Power and AI. AI Now report (27 September). Available at: https://ainowinstitute.org/publication/policy/compute-and-ai

viii Radu, R and Olliaro, E.(2023). Not child's play: protecting children data in humanitarian AI ecosystems. Humanitarian Law and Policy, ICRC: https://blogs.icrc.org/law-and-policy/2023/12/14/protecting-children-data-in-humanitarian-ai-ecosystems/

ix Biersteker, T. (2010). Global Governance. In *Routledge Handbook of Security Studies,* edited by Dunn-Cavelty, Myriam and Victor Mauer, pp. 439–51. London: Routledge.

x Jobin, A., Ienca, M. & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nat Mach Intell 1*, 389–399. https://doi.org/10.1038/s42256-019-0088-2; Ulnicane, I, Knight, W, Leach, T, Stahl, BC & Wanjiku, WG 2021, 'Framing governance for a contested emerging technology: insights from AI policy', *Policy and Society*, vol. 40, no. 2, pp. 158-177.

xi https://www.aretenews.com/us-china-talks-ai-risks-geneva/

xii A number of technical bodies, such as the ISO, IEC, IEEE, and the ITU are involved in AI standard-setting.

xiii Based on Radu, R. (2019). *Negotiating Internet Governance*. Oxford: Oxford University Press

xiv Feijóo, C. et al. (2020) Harnessing artificial intelligence (AI) to increase wellbeing for all: The case for a new technology diplomacy, Telecommunications Policy, Volume 44, Issue 6, https://doi.org/10.1016/j.telpol.2020.101988.

xv https://www.ebsco.com/resources/brics-wants-shape-global-ai-governance-too

xvi Huw, R., Hine, E., Taddeo, M, Floridi, L. (2024). Global AI governance: barriers and pathways forward, *International Affairs*, Volume 100, Issue 3, May 2024, Pages 1275–1286, https://doi.org/10.1093/ia/iiae073

IE UNIVERSITY
CENTER FOR THE GOVERNANCE OF CHANGE

**Written by:**
**Roxana Radu**, Associate Professor of Digital Technologies and Public Policy, University of Oxford.

**Recommended Citation:**
Radu, R., *The G20 and Global AI Governance*, IE CGC, July 2024.


Design: epoqstudio.com

For more information on the AI4Democracy initiative, visit:

ie.edu/cgc/research/ai4democracy