
THE FIRST GLOBAL AI TREATY: ANALYZING THE FRAMEWORK CONVENTION ON ARTIFICIAL INTELLIGENCE AND THE EU AI ACT

*Cheng-chi (Kirin) Chang**

This essay provides a comprehensive analysis of the Council of Europe's Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, adopted in May 2024. As the world's first legally binding international treaty on AI, the Convention aims to establish common standards for AI governance grounded in human rights, democratic values, and the rule of law. This essay examines the Convention's key provisions, comparing them with other regulatory frameworks, particularly the EU's Artificial Intelligence Act. It highlights the Convention's broad scope, lifecycle approach to AI governance, flexible implementation mechanisms, and emphasis on stakeholder engagement and international cooperation.

The analysis explores the Convention's strengths, including its global ambition, inclusive drafting process, and ethical foundations. However, it also critically assesses potential limitations, such as challenges in enforcement, possible regulatory fragmentation, and implementation hurdles in the face of political and technological complexities. This essay argues that while the Convention marks a crucial step towards coherent global AI governance, its effectiveness will ultimately depend on addressing these challenges and fostering a global culture of responsible AI development.

The essay concludes by offering recommendations for enhancing the Convention's impact, including developing supplementary protocols, strengthening monitoring mechanisms, and promoting ongoing

* Cheng-chi "Kirin" Chang (張正麒) is the Associate Director & Academic Fellow of the AI and the Future of Work Program at Emory University School of Law. I extend my appreciation to Dr. Ifeoma Ajunwa, J.D., LL.M., Ph.D., Yinn-ching Lu, Rachel Cohen, Yilin (Jenny) Lu, Nanfeng Li, Yenpo Tseng, Jeffrey Chang, Wolf (Chun-Ting) Cho, Zih-Ting You, Youyang Zhong, Ssu-Yuan (Iris) Yang, Arron Fang, Edison Li, Shijie Xu, and Yizhang (Yilia) Shen for their valuable insights and feedback on this article. Their contributions have significantly enhanced this work. I am grateful to Eli Goldstein, Michael Cerota, and the other editors of the *University of Illinois Law Review* for their diligent efforts in bringing this article to publication. Any errors or oversights are my sole responsibility. The views expressed in this are solely my own and do not represent those of any affiliated institutions.

international dialogue. It emphasizes the need for immediate next steps, such as refining the Convention through global stakeholder engagement, stress testing proposed measures, and expanding research to fill critical knowledge gaps. The Convention's success will be measured by its ability to guide the responsible development and deployment of AI technologies on a global scale, ensuring they serve to enhance rather than undermine human flourishing and societal well-being.

INTRODUCTION

As artificial intelligence systems grow more sophisticated by the day, from ChatGPT's viral success to breakthrough developments in robotics and autonomous vehicles, a critical governance vacuum looms. Who should write the rules governing AI's development and deployment? A handful of tech giants who control the technology? Individual nations pursuing their own interests? Or should we aspire to a truly global framework rooted in our shared values and human rights?

These questions lie at the heart of one of the most pressing challenges facing the international community today. As AI systems become increasingly sophisticated and ubiquitous, they promise unprecedented advancements across countless aspects of human society.¹ Yet, they also pose significant challenges to our fundamental rights, democratic institutions, and the rule of law.² Recognizing this urgent need for global governance, the Council of Europe has taken a groundbreaking step with the introduction of the world's first legally binding international treaty on AI - the Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law ("AI Framework Convention" or "Convention").³

Adopted on May 17, 2024, and opened for signature on September 5, 2024, this landmark Convention represents a watershed moment in the global effort to establish common standards for AI governance.⁴ It seeks to create a harmonized

1. Kristalina Georgieva, *AI Will Transform the Global Economy. Let's Make Sure It Benefits Humanity.*, IMF: BLOG (Jan. 14, 2024), <https://www.imf.org/en/Blogs/Articles/2024/01/14/ai-will-transform-the-global-economy-lets-make-sure-it-benefits-humanity> [https://perma.cc/YPW2-VXYL] ("We are on the brink of a technological revolution that could jumpstart productivity, boost global growth, and raise incomes around the world."). This highlights the promise of AI advancements and their potential benefits across various sectors in human society.

2. *See generally* DAVID LESLIE ET AL., HUMAN RIGHTS, DEMOCRACY, AND THE RULE OF LAW ASSURANCE FRAMEWORK FOR AI SYSTEMS: A PROPOSAL, ALAN TURING INST. (2021) (proposing a framework designed to assess the impact of AI on human rights and democracy, highlighting that AI systems can undermine democratic processes and individual rights if not properly regulated).

3. *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, C.E.T.S. No. 225, opened for signature Sept. 5, 2024 [hereinafter *AI Framework Convention*].

4. *Council of Europe Opens First Ever Global Treaty on AI for Signature*, COUNCIL OF EUROPE (Sept. 5, 2024), <https://www.coe.int/en/web/portal/-/council-of-europe-opens-first-ever-global-treaty-on-ai-for-signature> [https://perma.cc/3C2L-ZA3U].

approach among its signatories while respecting the diversity of legal systems and regulatory traditions across different jurisdictions.⁵

This essay provides a comprehensive analysis of the AI Framework Convention, examining its significance, key provisions, and potential to shape the future of global AI governance. By situating the Convention within the broader context of existing and emerging AI regulatory frameworks—most notably the European Union’s Artificial Intelligence Act (“EU AI Act”)⁶—we can better understand its unique contributions and limitations.

At its core, the AI Framework Convention aims to establish a common legal foundation for AI governance, grounded in the principles of human rights, democracy, and the rule of law. This essay argues that while the Convention marks a crucial step towards coherent global AI governance, its effectiveness will ultimately depend on addressing several key challenges. These include potential regulatory fragmentation, limitations in enforcement mechanisms, and the need to reconcile its broad principles with more detailed existing regulations.

The structure of this article is as follows: Part I provides essential background information on the development of the AI Framework Convention and the global context of AI governance efforts. Part II offers a detailed examination of the Convention’s key provisions, comparing and contrasting them with other relevant legal instruments, particularly the EU AI Act. Part III presents a critical analysis of the Convention’s potential impact, limitations, and future prospects, concluding with recommendations for enhancing its effectiveness and addressing remaining gaps in the international AI governance framework.

As we embark on this analysis, it is crucial to recognize that the AI Framework Convention represents not an endpoint, but rather a starting point in the ongoing global dialogue on AI governance. By critically engaging with its provisions and implications, we can contribute to the refinement and evolution of international legal approaches to ensuring that AI technologies serve to enhance, rather than undermine, our fundamental rights and democratic values.

I. BACKGROUND AND CONTEXT

The emergence of the Council of Europe’s AI Framework Convention must be understood within the broader context of global efforts to grapple with the societal implications of rapid advancements in AI technologies. Over the past decade, governments, international organizations, and civil society groups have increasingly recognized the need for coordinated action to address the potential risks and harness the benefits of AI systems across various domains of human activity.

5. *Id.*

6. Commission Regulation 2024/1689 of June 13, 2024, Laying Down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), 2024 O.J. (L 1689).

The Council of Europe, with its longstanding commitment to human rights, democracy, and the rule of law, has been at the forefront of these efforts. As early as 2019, the Committee of Ministers of the Council of Europe adopted a Declaration on the manipulative capabilities of algorithmic processes, highlighting the potential threats posed by AI-driven systems to democratic societies.⁷ This was followed by a series of recommendations and resolutions from various Council of Europe bodies, including the Parliamentary Assembly, which called for the development of core ethical principles to guide the deployment of AI systems.⁸

Parallel to these initiatives, other international organizations have also been working to establish guidelines and principles for AI governance. The Organisation for Economic Co-operation and Development (“OECD”) adopted its Recommendation on Artificial Intelligence in May 2019, setting out principles for the responsible stewardship of trustworthy AI.⁹ The United Nations Educational, Scientific, and Cultural Organization (“UNESCO”) followed suit with its Recommendation on the Ethics of Artificial Intelligence in November 2021, providing a comprehensive framework for ethical AI development and use.¹⁰

In the realm of binding legislation, the European Union has been leading the charge with its Artificial Intelligence Act (“EU AI Act”), first introduced in April 2021 and adopted on March 13, 2024.¹¹ The EU AI Act represents the world’s first comprehensive attempt to regulate AI systems through a risk-based approach, setting clear rules for high-risk applications while promoting innovation in the field.¹² This legislative initiative has set an important precedent and has significantly influenced the global discourse on AI regulation.

It is against this backdrop of proliferating soft law instruments and emerging hard law approaches that the Council of Europe embarked on the ambitious project of drafting a legally binding international convention on AI. The process began in earnest in May 2022, when the Council of Europe’s Committee on Artificial Intelligence (“CAI”) was tasked with elaborating a

7. *Council of Europe Adopts Declaration on Manipulative Capabilities of Algorithmic Processes*, DIGITAL WATCH OBSERVATORY (Feb. 13, 2019), <https://dig.watch/updates/council-europe-adopts-declaration-manipulative-capabilities-algorithmic-processes> [https://perma.cc/GW77-VN4F].

8. *Artificial Intelligence: Ensuring Respect for Democracy, Human Rights and the Rule of Law*, COUNCIL OF EUROPE, <https://pace.coe.int/en/pages/artificial-intelligence> (last visited Dec. 12, 2024) [https://perma.cc/BX9Q-PC3J].

9. *Recommendation of the Council on Artificial Intelligence*, OECD LEGAL INSTRUMENTS (2019) [hereinafter OECD], <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449> (last visited Dec. 12, 2024) [https://perma.cc/C6KR-Z7MU].

10. UNESCO, RECOMMENDATION ON THE ETHICS OF ARTIFICIAL INTELLIGENCE (2022).

11. Commission Regulation 2024/1689, *supra* note 6.

12. *EU AI Act: First Regulation on Artificial Intelligence*, EUROPEAN PARLIAMENT (June 18, 2024), <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence> [https://perma.cc/UJ7D-BK3G].

framework convention based on the Council's standards on human rights, democracy, and the rule of law.¹³

The drafting process was notable for its inclusivity and global reach. In addition to the Council of Europe's 46 member states, the negotiations involved participation from non-European observer states, including Argentina, Australia, Canada, Israel, Japan, Mexico, and the United States. The European Union, represented by the European Commission, also played an active role in the negotiations.¹⁴ Furthermore, a total of 68 civil society and industry representatives were involved as observers, ensuring a diverse range of perspectives in the Convention's development.¹⁵

This inclusive approach reflects a recognition of the inherently transnational nature of AI technologies and the need for global cooperation in addressing their challenges. As AI systems increasingly transcend national borders in their development, deployment, and impact, the limitations of purely domestic or regional regulatory approaches become apparent. The AI Framework Convention thus represents an attempt to forge a common legal framework that can serve as a foundation for more coherent and effective global AI governance.

The Convention's focus on human rights, democracy, and the rule of law distinguishes it from other AI governance initiatives that may prioritize economic or technical considerations.¹⁶ By anchoring the Convention in these fundamental values, the Council of Europe seeks to ensure that the development and use of AI technologies remain aligned with the principles that underpin democratic societies.¹⁷

It is important to note that the AI Framework Convention does not exist in isolation but is intended to complement and reinforce existing international human rights instruments. The Convention makes explicit reference to a wide range of global and regional human rights treaties, including the Universal

13. *COMMITTEE ON ARTIFICIAL INTELLIGENCE (CAI) ROADMAP: NEGOTIATIONS OF THE DRAFT [FRAMEWORK] CONVENTION*, COUNCIL OF EUROPE (2023).

14. EXPLANATORY REPORT TO THE COUNCIL OF EUROPE FRAMEWORK CONVENTION ON ARTIFICIAL INTELLIGENCE AND HUMAN RIGHTS, DEMOCRACY AND THE RULE OF LAW, COUNCIL OF EUROPE 1 (2024) [hereinafter EXPLANATORY REPORT] ("The Committee of Ministers also decided to allow for the inclusion in the negotiations of the European Union and interested non-European States sharing the values and aims of the Council of Europe – States from around the globe, namely Argentina, Australia, Canada, Costa Rica, the Holy See, Israel, Japan, Mexico, Peru, the United States of America and Uruguay, joined the process of negotiations in the CAI and participated in the elaboration of this Framework Convention as observer States.").

15. *Id.* ("A total of 68 civil society and industry representatives were involved in the CAI as observers, participating in the negotiations together with States and representatives of other international organisations.").

16. Bianca-Ioana Marcu, *The World's First Binding Treaty on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law: Regulation of AI in Broad Strokes*, FUTURE OF PRIVACY F. (June 20, 2024), <https://fpf.org/blog/the-worlds-first-binding-treaty-on-artificial-intelligence-human-rights-democracy-and-the-rule-of-law-regulation-of-ai-in-broad-strokes/> [https://perma.cc/2EQ5-N86D] ("The Framework Convention on AI Proposes a Risk-Based Approach and General Principles Focusing on Equality and Human Dignity.").

17. *The Framework Convention on Artificial Intelligence*, COUNCIL OF EUROPE, <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence> (last visited Dec. 12, 2024) [https://perma.cc/TH8B-U68B] ("It aims to ensure that activities within the lifecycle of artificial intelligence systems are fully consistent with human rights, democracy and the rule of law.").

Declaration of Human Rights, the International Covenants on Civil and Political Rights and on Economic, Social and Cultural Rights, and the European Convention on Human Rights, among others.¹⁸ This approach underscores the Convention's role in extending established human rights protections to the specific context of AI technologies.

As we delve deeper into the specific provisions of the AI Framework Convention in the following Section, it is crucial to keep this broader context in mind. The Convention represents not just a standalone legal instrument, but a key node in an evolving network of global governance initiatives aimed at ensuring that the transformative potential of AI technologies is harnessed in a manner that respects and promotes our fundamental rights and democratic values.

II. KEY PROVISIONS AND COMPARATIVE ANALYSIS

The AI Framework Convention comprises a preamble and 36 articles organized into eight chapters. This Section provides a detailed examination of the Convention's key provisions, comparing them with relevant aspects of the EU AI Act and other international instruments where appropriate.

A. *Scope and Definitions*

The Convention's scope, as defined in Article 3, is broad and encompasses "activities within the lifecycle of artificial intelligence systems that have the potential to interfere with human rights, democracy and the rule of law."¹⁹ This lifecycle approach is more comprehensive than that of the EU AI Act, which primarily focuses on the placing on the market, putting into service, and use of AI systems.²⁰ The Convention's approach recognizes that potential risks and impacts can arise at any stage of an AI system's lifecycle, from design and development to deployment and decommissioning.²¹

Notably, the Convention applies to both public authorities and private actors acting on their behalf, with a more flexible approach for other private actors.²² This is achieved through a declaration mechanism that allows Parties to specify how they intend to address risks and impacts arising from private sector

18. *AI Framework Convention*, *supra* note 3, at 2 ("Mindful of applicable international human rights instruments . . .").

19. *Id.* at 3.

20. *See generally AI Act Enters Into Force*, EUROPEAN COMMISSION (Aug. 1, 2024), https://commission.europa.eu/news/ai-act-enters-force-2024-08-01_en [<https://perma.cc/4ZGM-TPHN>]. The AI Act imposes specific obligations on providers of AI systems, particularly high-risk ones. These obligations include compliance with safety and ethical standards before an AI system can be placed on the market or put into service. Providers must undergo risk assessments, ensure data quality, and comply with transparency requirements.

21. EXPLANATORY REPORT, *supra* note 14, at 4 ("This reference to the lifecycle ensures a comprehensive approach towards addressing AI-related risks and adverse impacts on human rights, democracy and the rule of law by capturing all stages of activities relevant to artificial intelligence systems.")

22. *AI Framework Convention*, *supra* note 3, art. 3 (Article 3, paragraphs 1[a] and 1[b] of the Convention, which establish distinct approaches for public authorities and their representatives versus other private actors.).

AI activities.²³ This flexibility contrasts with the EU AI Act's more uniform application across public and private sectors. The EU AI Act applies to all providers of AI systems in the EU market and all users of AI systems located within the EU, regardless of whether they are public or private entities.²⁴

The Convention adopts the OECD's definition of an AI system, describing it as "a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that may influence physical or virtual environments."²⁵ This definition notably emphasizes the potential for AI systems to operate autonomously and adapt over time, reflecting a more dynamic understanding of AI capabilities. In contrast, the EU AI Act provides a more detailed and technically specific definition, which includes references to machine learning approaches, logic- and knowledge-based approaches, and statistical approaches.²⁶ This difference in definitions reflects the Framework Convention's broader, more flexible approach to encompassing evolving AI technologies, while the EU AI Act opts for a more precise, technically-grounded definition. The Convention's emphasis on autonomy and adaptiveness could potentially cover a wider range of AI systems, including those that may significantly change their behavior after deployment. This approach may prove more adaptable to rapid technological advancements but could also lead to broader interpretations of what constitutes an AI system under the Convention's purview.

B. General Obligations

Chapter II of the Convention sets out general obligations for Parties, focusing on the protection of human rights (Article 4) and the integrity of democratic processes and respect for the rule of law (Article 5).²⁷ These provisions reflect the Convention's grounding in fundamental democratic values, distinguishing it from more technically-oriented regulatory approaches.

23. *Id.* ("Each Party shall specify in a declaration submitted to the Secretary General of the Council of Europe at the time of signature, or when depositing its instrument of ratification, acceptance, approval or accession, how it intends to implement this obligation, either by applying the principles and obligations set forth in Chapters II to VI of this Convention to activities of private actors or by taking other appropriate measures to fulfil the obligation set out in this subparagraph.").

24. Commission Regulation 2024/1689, *supra* note 6, art. 2(1) (establishing that the regulation applies to both providers and users ["deployers"] of AI systems in the EU, regardless of whether they are located within or outside the EU. It covers both public and private entities implicitly by not making any distinction between them).

25. *AI Framework Convention*, *supra* note 3, at 3; OECD, *supra* note 9 ("AI system: An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.").

26. Commission Regulation 2024/1689, *supra* note 6, art. 3(1) ("'AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.").

27. *AI Framework Convention*, *supra* note 3, at 4.

Article 4 requires Parties to ensure that AI system activities are consistent with their human rights obligations under international and domestic law.²⁸ This broad approach allows for flexibility in implementation while reinforcing existing human rights frameworks in the context of AI.²⁹ The EU AI Act, while also emphasizing the protection of fundamental rights, takes a more prescriptive approach by defining specific prohibited AI practices and establishing a risk classification system for AI applications.³⁰

Article 5 addresses the potential impact of AI on democratic processes, requiring measures to protect against the undermining of democratic institutions and processes.³¹ This includes safeguarding the integrity of elections, public debate, and the formation of opinions.³² Such explicit protection of democratic processes is less prominent in the EU AI Act, which focuses more on specific use cases and risk categories.³³ The EU AI Act does, however, address some related concerns through its provisions on transparency and human oversight for high-risk AI systems.³⁴

C. Principles for AI Lifecycle Activities

Chapter III outlines key principles that should guide activities throughout the lifecycle of AI systems.³⁵ These include human dignity and individual autonomy, transparency and oversight, accountability and responsibility, equality and non-discrimination, privacy and personal data protection, reliability, and safe innovation.

The principle of human dignity and individual autonomy (Article 7) is not as prominently featured in the EU AI Act.³⁶ This principle underscores the Convention's human-centric approach, emphasizing the need to respect the inherent worth and agency of individuals in the face of increasingly autonomous AI systems.³⁷ While the EU AI Act does reference human dignity in its recitals, it does not elevate this concept to a central principle in the same way as the Convention.³⁸

28. *Id.*, art. 4.

29. *Id.*

30. Commission Regulation 2024/1689, *supra* note 6 (prescribing prohibited AI practices in Art. 5, establishing a risk classification system in Art. 6, and emphasizing fundamental rights protection throughout, e.g., Recital 46 mandating “common rules for high-risk AI systems” consistent with the Charter of Fundamental Rights).

31. *AI Framework Convention*, *supra* note 3, art. 5.

32. *Id.*

33. Commission Regulation 2024/1689, *supra* note 6 (classifying AI systems intended to influence elections or voting behavior as high-risk under Annex III, point 8[b]).

34. *Id.* (mandating transparency requirements for certain AI systems in Art. 50; requiring human oversight for high-risk AI systems in Art. 14, including systems used in democratic processes).

35. *AI Framework Convention*, *supra* note 3, at 4.

36. *AI Framework Convention*, *supra* note 3, art. 7.

37. *Id.*

38. Commission Regulation 2024/1689, *supra* note 6 (mentioning human dignity in Recitals 6 and 48 as a value to be protected, but not establishing it as a central principle or specific requirement in the operative articles of the Act).

The transparency and oversight principle (Article 8) aligns with similar requirements in the EU AI Act but extends beyond technical transparency to include the identification of AI-generated content.³⁹ This provision addresses growing concerns about the potential for AI to generate misleading or manipulative content that could undermine public discourse and democratic processes.⁴⁰ The EU AI Act includes transparency obligations for certain AI systems, such as chatbots, but does not explicitly address the identification of AI-generated content in the same manner as the Convention.⁴¹

The Convention's approach to equality and non-discrimination (Article 10) is more explicitly rights-based than the EU AI Act's. While both instruments aim to prevent discriminatory outcomes from AI systems, the Convention frames this principle more directly in terms of human rights obligations.⁴² The EU AI Act addresses discrimination primarily through its risk assessment framework and requirements for high-risk AI systems.⁴³

D. Remedies and Procedural Safeguards

Chapter IV of the Convention focuses on remedies (Article 14) and procedural safeguards (Article 15), emphasizing the importance of effective redress mechanisms for individuals affected by AI systems.⁴⁴ These provisions require Parties to ensure the availability of accessible and effective remedies for human rights violations resulting from AI activities.

Notably, Article 14 mandates that relevant information about AI systems that significantly affect human rights be documented and made available to affected persons.⁴⁵ This transparency requirement goes beyond similar provisions in the EU AI Act, potentially facilitating more effective contestation of AI-driven decisions.⁴⁶ While the EU AI Act includes requirements for documentation and record-keeping for high-risk AI systems, it does not explicitly require this information to be made available to affected individuals in the same manner as the Convention.⁴⁷

39. *AI Framework Convention*, *supra* note 3, art. 8.

40. *Id.*

41. Commission Regulation 2024/1689, *supra* note 6 (mandating transparency for AI systems interacting with humans in Art. 50[1], and requiring providers to ensure AI-generated content is detectable in Art. 50[2], but not explicitly addressing all AI-generated content identification in the same comprehensive manner).

42. *AI Framework Convention*, *supra* note 3, art. 10.

43. Commission Regulation 2024/1689, *supra* note 6 (addressing discrimination through the risk assessment framework in Art. 9, data quality requirements in Art. 10, and classifying AI systems with potential discriminatory impacts as high-risk in Annex III, particularly in employment, education, and access to services contexts).

44. *AI Framework Convention*, *supra* note 3, at 5.

45. *Id.*, art. 14.

46. *Id.*

47. Commission Regulation, 2024/1689, *supra* note 6 (requiring documentation and record-keeping for high-risk AI systems in Arts. 11 and 12, but primarily for regulatory compliance rather than individual access; providing limited rights for affected persons to obtain explanations in Art. 86, but not mandating comprehensive access to system documentation).

Article 15 introduces procedural safeguards, including the requirement to notify individuals when they are interacting with an AI system rather than a human.⁴⁸ This provision addresses concerns about the potential for AI to manipulate or deceive users, an issue not explicitly covered in the EU AI Act in the same way.⁴⁹ The EU AI Act does require transparency for certain AI systems that interact with humans, but the Convention's approach is more broadly applicable across different types of AI systems.⁵⁰

E. Risk and Impact Management

Chapter V establishes requirements for risk and impact management frameworks (Article 16).⁵¹ This approach shares similarities with the EU AI Act's risk-based framework but adopts a more flexible and context-sensitive approach. The Convention requires Parties to adopt measures for identifying, assessing, preventing, and mitigating risks posed by AI systems, taking into account the severity and probability of potential impacts.⁵² A notable feature of the Convention's approach is its emphasis on stakeholder engagement, particularly the involvement of those who may be affected by AI systems. Article 16(2)(c) explicitly requires Parties to "consider, where appropriate, the perspectives of relevant stakeholders, in particular persons whose rights may be impacted" when implementing risk management measures.⁵³ This provision underscores the Convention's commitment to a participatory approach in AI governance, recognizing the importance of incorporating diverse perspectives in risk assessment and mitigation strategies.

While both the Convention and the EU AI Act emphasize the importance of risk management, their approaches differ in significant ways. The EU AI Act establishes a tiered risk classification system, with specific requirements for each risk level.⁵⁴ In contrast, the Convention provides a more general framework for risk assessment and management, allowing Parties greater flexibility in how they implement these requirements.⁵⁵ The EU AI Act does include provisions for stakeholder consultation, particularly in the context of developing standards and codes of conduct, but it does not explicitly mandate consideration of affected persons' perspectives in the risk management process to the same extent as the Convention.⁵⁶

48. *AI Framework Convention*, *supra* note 3, art. 15.

49. *Id.*

50. Commission Regulation 2024/1689, *supra* note 6 (mandating transparency for specific AI systems interacting with humans in Art. 50[1] and for certain AI-generated content in Art. 50[2], but limiting these requirements to particular use cases rather than applying them broadly across all AI systems).

51. *AI Framework Convention*, *supra* note 3, at 6.

52. *Id.*, art. 16.

53. *Id.*, art. 16(2)(c).

54. Commission Regulation 2024/1689, *supra* note 6 (establishing a tiered risk classification system in Art. 6, with specific requirements for high-risk AI systems in Chapter III, Section 2, and different obligations for general-purpose AI models in Chapter V, while prohibiting certain AI practices in Art. 5).

55. *AI Framework Convention*, *supra* note 3, art. 16.

56. *Id.*

Notably, Article 16 of the Convention also requires Parties to assess the need for moratoria, bans, or other measures for certain AI uses deemed incompatible with human rights, democracy, and the rule of law.⁵⁷ This provision provides a mechanism for addressing high-risk AI applications that may not be adequately covered by existing regulatory frameworks. The EU AI Act includes a list of prohibited AI practices but does not explicitly call for ongoing assessment of the need for bans or moratoria in the same way as the Convention.⁵⁸

F. Implementation and Follow-up Mechanism

Chapters VI and VII outline the Convention's implementation requirements and follow-up mechanism. These include provisions on non-discrimination (Article 17), rights of persons with disabilities and children (Article 18), public consultation (Article 19), and digital literacy (Article 20).⁵⁹

The Convention establishes a Conference of the Parties (Article 23) to monitor implementation and facilitate information exchange.⁶⁰ Additionally, Article 26 requires Parties to establish or designate effective oversight mechanisms to ensure compliance with the Convention's obligations.⁶¹ This provision strengthens the Convention's implementation framework by mandating national-level accountability. In contrast, the EU AI Act creates a more structured governance system, including a European Artificial Intelligence Board and national competent authorities for oversight and enforcement.⁶² The Convention's approach offers greater flexibility but may result in varied implementation across Parties.

Article 25 on international cooperation emphasizes the need for Parties to exchange information and collaborate on preventing and mitigating risks to human rights, democracy, and the rule of law.⁶³ This explicit focus on international cooperation distinguishes the Convention from more regionally-focused instruments like the EU AI Act.⁶⁴ While the EU AI Act includes provisions for cooperation among EU member states and with third countries, it does not have the same global scope as the Convention.

In summary, while the AI Framework Convention shares some common ground with the EU AI Act, particularly in its risk-based approach and emphasis

57. *Id.*

58. Commission Regulation 2024/1689, *supra* note 6 (listing prohibited AI practices in Art. 5, with Art. 112[1] requiring annual assessment of the need to amend this list, but not explicitly mandating ongoing assessment of potential bans or moratoria beyond the existing framework).

59. *AI Framework Convention*, *supra* note 3, at 6–9.

60. *Id.*, art. 23.

61. *Id.*, art. 26.

62. Commission Regulation 2024/1689, *supra* note 6 (establishing a structured governance framework including the European Artificial Intelligence Board in Art. 65 and designating national competent authorities for oversight and enforcement in Art. 70, creating a more rigid regulatory structure compared to flexible approaches).

63. *AI Framework Convention*, *supra* note 3, art. 25.

64. *Id.*

on fundamental rights, it distinguishes itself through its global scope, explicit focus on democratic values, and more flexible implementation framework. The Convention's broader lifecycle approach and emphasis on international cooperation positions it as a potentially significant instrument for shaping global AI governance norms. However, its effectiveness in practice will depend on how it is implemented and enforced by its signatories, a challenge we will explore further in the next Section.

III. CRITICAL ANALYSIS AND FUTURE PROSPECTS

The AI Framework Convention marks a significant milestone in global AI governance, yet its effectiveness hinges on both its strengths and limitations. Its global ambition and inclusive drafting process distinguish it from regional initiatives like the EU AI Act, fostering broader international consensus on AI governance principles. By grounding the Convention in human rights, democratic values, and the rule of law, it provides a robust ethical foundation for AI governance, reinforcing the alignment of technological advancement with fundamental societal values.

However, the Convention faces significant challenges. Despite its "legally binding" status, it lacks strong enforcement mechanisms which contrasts sharply with the EU AI Act's robust framework.⁶⁵ The Convention's broad, principle-based provisions could lead to divergent interpretations and implementations among Parties, potentially undermining harmonized international standards. Its flexible approach to private sector regulation could result in regulatory fragmentation, complicating compliance for multinational AI developers.

Implementing the Convention requires unprecedented international cooperation, which may be challenging given political polarization and geopolitical tensions.⁶⁶ Technological hurdles also exist, such as enforcing risk assessment requirements for different training run thresholds, which may require detecting and quantifying global computing power usage.⁶⁷ Governance challenges include determining safety assurance requirements for potentially catastrophic technologies and measuring such assurance.⁶⁸

65. Javier Espinoza & Madhumita Murgia, *US, Britain and Brussels to Sign Agreement on AI Standards*, (Sept. 5, 2024), <https://www.ft.com/content/4052e7fe-7b8a-4c42-baa2-b608ba858df5> [<https://perma.cc/8GD3-JDEZ>] ("While the treaty is billed as 'legally enforceable,' critics have pointed out that it has no sanctions such as fines. Compliance is measured primarily through monitoring, which is a relatively weak form of enforcement.")

66. See generally Richard Danzig, *Technology Roulette: Managing Loss of Control as Many Militaries Pursue Technological Superiority* (May 30, 2018), <https://www.cnas.org/publications/reports/technology-roulette> [<https://perma.cc/8CDZ-BD8D>]; Christian Ruhl & Stephen Clare, *Great Power Competition and Transformative Technologies Report*, CTR. NEW AM. SEC. (JAN. 31, 2024), <https://www.founderspledge.com/research/great-power-competition-and-transformative-technologies-report> [<https://perma.cc/8HDF-HYVQ>].

67. GIRISH SASTRY ET AL., *COMPUTING POWER AND THE GOVERNANCE OF ARTIFICIAL INTELLIGENCE*, ARXIVLABS (2024).

68. See generally JOHN DOWNER, *RATIONAL ACCIDENTS: RECKONING WITH CATASTROPHIC TECHNOLOGIES* (2024).

Despite these challenges, the seriousness of risks and potential benefits create incentives for cooperation. Public wariness of advanced AI and support for risk reduction efforts provide domestic political incentives for international action.⁶⁹ Precedents for international cooperation on AI and technological risks exist, including shared interests between the United States and China on issues like autonomous weapons oversight.⁷⁰

To enhance the Convention's impact and address its limitations, several strategies could be employed. Developing supplementary protocols to address specific AI applications or emerging challenges could help fill gaps in the current framework and provide more detailed guidance to Parties. This approach would allow the Convention to maintain its broad principles while addressing the unique challenges posed by contentious AI applications such as facial recognition or autonomous weapons systems.

Strengthening the monitoring and enforcement mechanisms, perhaps by enhancing the powers and resources of the Conference of Parties or establishing a dedicated monitoring body, could improve oversight of the Convention's implementation. This would address the current weakness in enforcement and bring the Convention closer to the robust framework of the EU AI Act.

Fostering ongoing international dialogue among Parties, as well as with nonparty states and relevant international organizations, could help build consensus on the interpretation and implementation of the Convention's provisions. Creating detailed implementation guidelines or best practices could address ambiguity in some of the Convention's provisions and promote more consistent application across different jurisdictions.

Closer coordination with other international AI governance efforts, such as those led by the UN, OECD, and regional bodies, could help reduce regulatory fragmentation and promote a more coherent global approach to AI governance. Developing programs to support capacity building, particularly for developing countries, could ensure more equitable participation in the Convention's implementation and the broader global AI governance landscape.

Continuing to engage a diverse range of stakeholders, including civil society organizations, industry representatives, and academic experts, in the Convention's ongoing development and implementation could help ensure its relevance and effectiveness. This multistakeholder approach will be crucial in addressing the complex and rapidly evolving challenges posed by AI technologies.

Immediate next steps could include refining the proposed Framework Convention through engagement with global stakeholders, stress testing proposed measures for effectiveness under challenging scenarios and encouraging ongoing multilateral negotiations to adopt proposed measures.

69. DUNCAN CASS-BEGGS, STEPHEN CLARE, DAWN DIMOWO & ZAHEED KARA, *FRAMEWORK CONVENTION ON GLOBAL AI CHALLENGES*, CTR. INT'L GOVERNANCE INNOVATION 19 (2024).

70. *Id.*; See Ryan Hass & Colin Kahl, *Laying the Groundwork for US-China AI Dialogue*, BROOKINGS (Apr. 5, 2024), <https://www.brookings.edu/articles/laying-the-groundwork-for-us-china-ai-dialogue/> [https://perma.cc/NX8C-M7Y7].

Expanding research and policy development efforts to fill critical knowledge gaps and initiating processes to draft and adopt international agreements with necessary measures and institutions are also crucial.

The AI Framework Convention offers a valuable foundation for global AI governance. Its success will ultimately depend on widespread ratification, meaningful implementation, and addressing current limitations. As AI technologies evolve, sustained commitment from signatories, ongoing multistakeholder engagement, and willingness to adapt will be crucial in realizing the Convention's potential to guide responsible AI development on a global scale.

By fostering a shared ethical foundation and promoting international cooperation, the Convention can play a vital role in ensuring that AI technologies enhance rather than undermine human flourishing and societal well-being. The journey towards effective global AI governance has only just begun, and the Convention marks an important first step on this critical path. As we navigate the unprecedented challenges and opportunities presented by AI technologies, such a shared ethical foundation will be indispensable in shaping a future where AI serves humanity's best interests.

CONCLUSIONS

The AI Framework Convention represents a significant milestone in the global effort to establish common standards for AI governance, anchored in the principles of human rights, democracy, and the rule of law. As the world's first legally binding international treaty on AI, it offers a flexible and inclusive approach to addressing the complex challenges posed by rapidly evolving AI technologies. However, the Convention's effectiveness will ultimately depend on widespread ratification, meaningful implementation, and the ability to address its current limitations.

The Convention's strengths lie in its global scope, comprehensive lifecycle approach, and grounding in fundamental democratic values. Yet, it faces challenges in terms of enforcement mechanisms, potential regulatory fragmentation, and reconciliation with more detailed existing regulations. Moving forward, the success of the AI Framework Convention will hinge on sustained international cooperation, ongoing stakeholder engagement, and the development of supplementary protocols and implementation guidelines. As AI continues to transform our societies, the Convention provides a crucial foundation for fostering a global culture of responsible AI development that respects human rights, upholds democratic values, and strengthens the rule of law.