Conceptualizing Global Governance of AI

Maral Niazi

Fall 2023 cohort

About the Hub

The Digital Policy Hub at CIGI is a collaborative space for emerging scholars and innovative thinkers from the social, natural and applied sciences. It provides opportunities for undergraduate and graduate students and post-doctoral and visiting fellows to share and develop research on the rapid evolution and governance of transformative technologies. The Hub is founded on transdisciplinary approaches that seek to increase understanding of the socio-economic and technological impacts of digitalization and improve the quality and relevance of related research. Core research areas include data, economy and society; artificial intelligence; outer space; digitalization, security and democracy; and the environment and natural resources.

The Digital Policy Hub working papers are the product of research related to the Hub's identified themes prepared by participants during their fellowship.

Partners

Thank you to Mitacs for its partnership and support of Digital Policy Hub fellows through the Accelerate program. We would also like to acknowledge the many universities, governments and private sector partners for their involvement allowing CIGI to offer this holistic research environment.

About CIGI

The Centre for International Governance Innovation (CIGI) is an independent, non-partisan think tank whose peer-reviewed research and trusted analysis influence policy makers to innovate. Our global network of multidisciplinary researchers and strategic partnerships provide policy solutions for the digital era with one goal: to improve people's lives everywhere. Headquartered in Waterloo, Canada, CIGI has received support from the Government of Canada, the Government of Ontario and founder Jim Balsillie.

<div style="border: 1px solid green;">

Key Points

- Merging artificial intelligence (AI) and global governance, global AI governance focuses on defining terms to deepen understanding, promote collaboration and create informed policies. It emphasizes multi-stakeholder and multi-level cooperation in managing AI's global impacts.

- AI's societal impacts are broad, offering exceptional benefits while carrying unintended risks. Its rise poses geopolitical challenges, affecting transparency, privacy and power dynamics in both democratic and non-democratic states.

- Empirical and normative research is essential in forming global AI governance, guiding ethical values and legal practices for ethical data use and unbiased algorithm development. Empirical research provides verifiable knowledge through data and experiences, highlighting regime complexities in an anarchic system of global governance, meaning a system with no central authority. Normative research examines values and norms, assessing AI systems' trustworthiness and ethical compliance.

- Multilateral cooperation in global AI governance involves collaborative efforts among numerous actors to establish universally accepted norms and policies for AI. Key attributes include the following: collaborative research, uniting diverse stakeholders to align policies with global ethical standards; flexible governance, adaptable to evolving AI and diverse needs; common standards for responsible AI development; harmonized regulation for unified legal and ethical standards; and universal compliance for global consistency in AI practices.

- An institutional framework for global AI governance should incorporate lessons from international organizations such as the International Atomic Energy Agency (IAEA) and the European Organization for Nuclear Research/Conseil européen pour la recherche nucléaire (CERN) to guide AI's ethical development and deployment within and beyond national borders.

</div>

Note: This working paper was completed just prior to the release of the interim report of the UN Secretary-General's AI Advisory Body, *Governing AI for Humanity*, in December 2023.

Introduction

The advancement of AI has pressed policy makers and AI developers to focus on ethical AI that reduces algorithm and data biases. Because the top hegemonic powers are leading the race for AI supremacy, AI will have implications for global power dynamics. AI's integration into the public sphere presents technological revolutionary potential in sectors such as education (Cummings et al. 2018; Ulnicane et al. 2022). AI regulation remains in its infant stage. Notably, the UN Education, Scientific and Cultural Organization (UNESCO) has just begun to develop ethical guidelines, the European Union has recently enacted the comprehensive AI Act, and the G7 is still working to advocate for global AI standards. These initiatives aim to shape the future of AI governance, power distribution and economic dynamics (Erman and Furendal 2022; Tallberg et al. 2023). The dangers posed by

AI, including enhanced pathogens, underscore the need for a global governance framework that extends beyond regulation. This framework requires extensive collaboration among a diverse group of stakeholders and a deep understanding of AI's foundational algorithms and data to ensure ethical, rules-based development.

To examine the multilateral approach to the global governance of AI, it is pivotal to define AI, global governance and global AI governance. Next, it is key to identify the main issues and impacts of AI. Furthermore, empirical and normative research in global AI governance will help bridge the field of AI and global governance, and address the misconceptions associated with AI governance. Additionally, it is important to explain why a multilateral approach is most suitable for global AI governance in utilizing AI's benefits and mitigating its risks. Finally, as a recommendation, examining the existing models of multilateral institutions is imperative in conceptualizing the foundation of a multilateral global AI governance structure.

Definitions

Global AI governance is a brand-new term that is increasingly being used in the discussion of regulating AI. As the two fields of AI and global governance are merging to give birth to a new field of study, defining its terms is fundamental in establishing scholarly discourse (Jobin, Ienca and Vayena 2019). These definitions will facilitate dissecting the multi-faceted nature of AI governance and gaining a deeper understanding of its impacts and issues. They will also enable a structured approach to examining AI, promoting collaboration and informed policy making on a global scale (Floridi 2019; Floridi and Cowls 2019).

AI Definition

John McCarthy defines AI as "the science and engineering of making intelligent machines" (quoted in Hamet and Tremblay 2017). AI functions primarily through algorithms and data. Algorithms enable AI systems to perform tasks by providing them with rules to process information. At the same time, data is used to train these algorithms, allowing AI systems to adapt and make decisions (Russell and Norvig 2016). At the heart of AI is machine learning, and within that deep learning, two subsets of AI where machines learn from data and make decisions with minimal human intervention (Janiesch, Zschech and Heinrich 2021).

Global Governance Definition

Global governance is a set of collective agreements and entities based on common norms and procedures regulating transnational interactions with no centralized global authority (Weiss 2016; Rosenau and Czempiel 1992). In this anarchic model of governance, states, empowered by sovereignty, and non-state actors, such as international organizations, interact through diplomacy to solve complex global issues to maintain international peace and security (Rhodes 1996; Rosenau and Czempiel 1992).

Global AI Governance Definition

Global AI governance is a multi-stakeholder regulatory framework that facilitates collective actions and transnational cooperation to ensure responsible AI development. Previous scholars, such as Luciano Floridi (2019) and Josh Cowls (2019), have defined global governance as the development and administration of legal standards and ethical norms and policies for AI at the global level. The issue is that definitions like Floridi's do not take into account the importance of multilateral cooperation. Multilateral cooperation completes global AI governance by bringing together diverse international perspectives, thereby ensuring comprehensive management of AI's global challenges.

AI's Impacts and Issues

AI and policy experts have yet to fully understand the scope of AI; as a result, AI's societal impacts are arguably broad ranging. However, there are certain areas where strong evidence reveals the consequences of AI technologies. AI offers significant benefits across society (Azoulay 2018; Jobin, Ienca and Vayena 2023; Mäntymäki et al. 2022). In health care, AI algorithms have revolutionized diagnostics and patient treatment; for instance, AI-driven analysis in radiology significantly improves the detection and interpretation of medical images. Deep learning advancements provide more accurate medical diagnoses and data analysis in DNA and RNA sequencing (Hosny et al. 2018). AI also plays a critical role in climate change research; machine learning models guide more accurate climate predictions and provide "transformational solutions" (UN Framework Convention on Climate Change 2023, 1).

However — like the internet — AI carries notable issues. One prominent concern is where AI harms individual rights and freedoms in employment, personal data protection and algorithm biases. Rights abuses raise serious ethical concerns when authoritarian regimes use AI to increase their surveillance capabilities on their citizens, causing various forms of social oppression (Bostrom and Yudkowsky 2014; Ho et al. 2023; Sepasspour 2023). China's government has integrated AI into its governance mechanisms, which can potentially suppress opposition and reinforce state control (Zeng 2022). Depending on who governs it, AI systems can infringe on privacy rights and freedom of expression, causing discrimination and bias (Donahoe and Metzger 2019). In particular, examining the societal impacts of AI on the economy and geopolitics are essential components of studying global AI governance (Butcher and Beridze 2019; Liebig et al. 2022). The rise of AI technologies presents significant geopolitical challenges, particularly in the differing approaches to AI governance and use between democratic and non-democratic states. On the one hand, democratic states often emphasize transparency, privacy and individual rights in AI development; however, the rise of AI in capitalist states such as the United States can potentially concentrate power in the hands of corporate entities, leading to a form of corporate dictatorship (Zuboff 2019). On the other hand, non-democratic states may leverage AI for mass surveillance and social control, exemplified by China's use of AI in public surveillance systems (Carnegie Endowment for International Peace 2019).

Empirical and Normative Understanding of Global AI Governance

Even though global AI governance is a concern for policy makers, little empirical and normative research has been done regarding it. Empirical and normative research should form the foundation for creating a global AI governance framework. Ethical values and legal practices should guide the agreements that come with it. A global AI governance framework should address challenges related to the use of data and the development of algorithms within and beyond national boundaries. Empirical and normative research are both able to help identify how this framework can address AI technology's ever-evolving landscape (Hoffmann-Riem 2019; UNESCO 2019). This dual approach can help formulate effective policies that make the governing system technically sustainable and equitable.

Empirical Understanding of Global AI Governance

Empirical research is a method of knowledge acquisition or experimentation from verifiable experiences, data and facts. One of the most critical avenues to form a comprehensive perspective of the global governance of AI is to study it through an empirical lens that highlights regime complexities. This empirical lens reflects what global AI governance is supposed to be — a network of partially overlapping and diverse governance mechanisms without a centralized authority (Erman and Furendal 2022; Tallberg et al. 2023).

An empirical approach through international relations and global governance theories will underscore significant elements impacting the foundation where the entire global AI governance framework stands. International relations theories are important for a more comprehensive empirical study of global AI governance as they provide impactful foundations for understanding the complex interplay of global power dynamics in shaping AI common standards. This framework should consider power dynamics, actors and conventions such as hard and soft laws, which help make it function. Actors influence AI governance, each playing a pivotal role in shaping the governance framework (Erman and Furendal 2022; Tallberg et al. 2023).

Western Approaches

The European Union's General Data Protection Regulation (GDPR) exemplifies a centralized approach by binding EU nations to set standards for data privacy that have global implications.[1] The GDPR parallels global AI governance in its comprehensive approach to regulating data (ibid.). The GDPR aims to address issues of data privacy that focus on AI systems processing personal data. Just as GDPR emphasizes individual rights of personal data, global AI governance also focuses on user consent and algorithmic transparency (Russell and Norvig 2016). The GDPR's global influence, affecting AI

---

1   See https://gdpr-info.eu/; https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence; https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai.

companies worldwide, mirrors the cross-border implications of AI governance, showcasing how regulatory frameworks can set international precedents and standards (Voigt and von dem Bussche 2017).[2] Furthermore, the GDPR's enforcement mechanisms for non-compliance are akin to the accountability measures that are essential in global AI governance, ensuring adherence to established ethical and legal standards.[3]

The US National AI Initiative underscores the role of national strategies in steering AI research and development (R&D) (National Science Foundation [NSF] 2021; National Science and Technology Council [NSTC] 2023). The US approach to AI governance adopts a more decentralized model than the European Union, which emphasizes the role of market forces in guiding AI R&D. Based on the National Artificial Intelligence Initiative Act of 2020, the United States prioritizes flexibility and responsiveness to technological advancements, focusing on R&D, ethical considerations and workforce development (NSTC 2023).[4] The downsides could lead to challenges in establishing uniform ethical standards and comprehensive regulatory oversight. The US decentralization culture pushes for quicker deployment of AI to the general public. However, unlike the European Union, it does not focus on universal legislation that would mitigate any impacts from faster integration into society.

Non-Western Approach

China showcased a different governance model with its centralized AI strategy under the New Generation Artificial Intelligence Development Plan, prioritizing state-led directives and large-scale investments (Webster et al. 2021). China's divergence highlights a fundamental tension in global AI governance: the EU method underscores precaution and control, while the US model champions technological advancement and economic competitiveness, each shaping AI's global trajectory differently. China's state-led, centralized AI strategy presents a third paradigm that combines governmental control with aggressive investment in AI.

China's plan challenges the liberal market-centric approach of the United States and the regulatory stringency of the European Union (Webster et al. 2017). These varied national approaches indicate how global AI governance has been shaped by differing political and economic contexts in the past few years, offering guidance and insights for other countries and regions of the world to set out their AI governance models at national and transnational levels.

Non-state Actors

At the international level, examples of international organizations that have an AI governance agenda include the Organisation for Economic Co-operation and Development (OECD) with its AI Principles and UNESCO with its ethics recommendations (OECD 2019; UNESCO 2019). Through these frameworks, these international organizations aim to conform to standards of national policies and corporate practices. The OECD and UNESCO's

---

2   See https://gdpr-info.eu/.

3   See https://gdpr-info.eu/; https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence; https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai.

4   See https://new.nsf.gov/focus-areas/artificial-intelligence.

plans to push global AI governance forward highlight the strategic role of international organizations in setting global norms in AI governance.

In addition to international organizations, the significant contributions of other non-state actors, particularly corporations and civil societies, should be addressed. For instance, the Partnership on AI (PAI), a collective of technology companies and non-profits, influences ethical AI development, notably through publications, such as the "Safety-Critical AI" white paper, which provides comprehensive frameworks for creating trustworthy AI systems (PAI 2020). The PAI's role in ethical AI development exemplifies how non-state actors, such as corporations and civil society groups, alongside international organizations, create a multi-stakeholder approach in global AI governance (ibid.).

Normative Understanding of Global AI Governance

Normative research involves the study of values and norms, focusing on what ought to be rather than what is, often addressing ethical or moral standards. Normative research falls into three categories: distinguishing how the process of governance development functions; explaining the content of the governance itself; and assessing AI systems governed under these frameworks. The multi-faceted approach of normative research enables a comprehensive understanding of the effectiveness and ethical alignment of AI governance (Erman and Furendal 2022; Tallberg et al. 2023).

The process of developing AI governance is crucial in determining its functionality. A good governance system should be developed through inclusive and democratic means. This includes wide-ranging stakeholder engagement, where not only government bodies and technology companies have a say in policy formulation, but also civil society, academia and under-represented groups (Erdélyi and Goldsmith 2022). For instance, the European Union's approach to formulating the AI Act involved broad public consultations, ensuring diverse perspectives were considered.[5] Such inclusive processes enhance the governance system's credibility and ensure that it addresses the needs and concerns of a wide spectrum of society.

The content of AI governance — encompassing principles, norms and outcomes — is equally critical. It must align with global ethical standards and address key concerns such as fairness, transparency and accountability. The OECD's "Ethical Guidelines for Artificial Intelligence" presents a good governance framework due to its comprehensive and pragmatic approach, as it encourages AI actors to commit to transparency and responsible disclosure regarding AI systems. Providing meaningful information will help enhance a general understanding of AI systems and mitigate its risks (OECD 2019; 2023). The effectiveness of these principles lies in their ability to guide AI development and use toward ethical and socially beneficial outcomes. However, the challenge often lies in translating high-level principles into actionable policies that can be consistently applied across different contexts and jurisdictions.

The normative assessment of global AI governance requires an evaluation of the process of governance development, the content of the governance itself and the quality of AI systems it produces. Due to evidence of AI biases, a powerful assessment

---

5   See https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence.

metric would be AI systems' trustworthiness and ethical compliance (P. S. 2023). The development of ethical AI is exemplified in projects such as IBM's AI Fairness 360, which provides tools to detect and mitigate bias in AI models (IBM 2018). Trustworthy AI systems are those that not only comply with technical and ethical standards but also align with societal values and contribute positively to societal challenges. Ethical AI systems require that they be developed with respect for user privacy and consent, aligning with the values and rights of individuals (Cath 2018). Without rigorous normative assessments, AI biases and the lack of regulatory mechanisms to combat them will unfairly affect people around the world.

Multilateral Cooperation

Multilateral cooperation, defined as collaborative efforts among multiple countries or international actors, is pivotal in cultivating global AI governance by creating universally accepted norms, standards and policies that address the complexities and ethical challenges of AI technology across borders (Schmitt 2021). Multilateralism has five major attributes that define it as a whole:

- **Collaborative research** is the joint efforts of diverse stakeholders to collectively study, develop and refine policies, ensuring they are aligned with global ethical standards and societal needs (Vinuesa et al. 2020).

- **Flexible governance** is a regulatory framework and ethical guidelines that are adaptable and responsive to the evolving nature of governance mechanisms and diverse global needs (Cath 2018).

- **Common standards** are universally agreed-upon guidelines and ethical norms that govern the development of regulation, ensuring their responsible and equitable application across nations (Jobin, Ienca and Vayena 2019).

- **Harmonized regulation** is the establishment of unified and consistent legal and ethical standards across different countries, ensuring a coherent approach to unpredictable changes in global governance (Cath 2018).

- **Universal compliance** is the uniform adherence of countries and organizations to internationally agreed-upon standards and regulations, ensuring a consistent and ethical demeanour globally (Jobin, Ienca and Vayena 2019).

Collaborative research is fundamental in establishing a framework for governance and setting up multilateral standards and regulations at the global level. Collaborative research brings together diverse perspectives, which is crucial for developing inclusive standards that promote trust among nations. Engaging in joint research efforts helps prevent regulatory conflicts from isolated national strategies. Furthermore, it gives developing countries a voice in shaping global policies. Without collaborative research, multilateral cooperation cannot create the other four attributes it requires to function (Lopez-Claros, Dahl and Groff 2020; Mäntymäki et al. 2022).

The current multilateral global AI governance framework is in its infancy as AI outpaces international regulatory framework development. It stems from challenges in establishing flexible governance, which should be responsive to the

evolving AI landscape. Achieving common standards, universal compliance and harmonized regulation are complex, requiring consensus among nations (Miailhe and Lannquist 2020). Furthermore, effective collaborative research necessitates extensive cooperation across countries. Despite these challenges, a global AI governance framework is necessary to complement regional and national strategies, promoting more equitable access to AI within fragmented regulation networks.

Recommendation: Examine the Existing Mechanisms of Multilateral Institutions

As the landscape of global AI governance evolves, the role of established multilateral institutions is becoming increasingly crucial. With their diverse international mandates, some of the existing institutions hold the potential to offer constructive models that shape the development and application of AI in ways that align with global security, ethics and human welfare (Hendrix 2023). The following are recommendations for creating one new multilateral institution for the governance of AI inspired by existing institutions such as CERN.

- **IAEA for AI non-proliferation agreements:** A model similar to the IAEA can help set up non-proliferation AI agreements. Including this model as one of the main foundations of the global AI governance framework can help govern AI technologies such as autonomous weaponry. This model aims to develop a comprehensive AI non-proliferation treaty similar to nuclear non-proliferation agreements to prevent the spread of AI technologies that pose significant threats to humanity and global security (Nichols 2023). Such treaties can help increase transparency and trust in the AI industry, as developers and governments would be required to report on their AI R&D activities, leading to greater accountability for the potential misuse of AI, as violators of the treaty would face sanctions and penalties. Despite the criticism of the IAEA on issues of espionage from international auditors, collaborative research can investigate accountability mechanisms that do not violate the country's sovereignty.

- **CERN for AI in service of humanity:** Embedding a model similar to CERN within the global AI governance framework can help initiate programs that leverage AI for humanitarian causes, such as managing climate change, disease control and sustainable development. An open access to information approach could increase the use of AI for humanitarian purposes in lesser-developed regions as it promotes collaborative and transparent research in the AI world. It also boosts AI development and innovation, as a CERN-like model would create a fertile ground for new AI techniques and applications in the service of humanity, regardless of political and economic discrimination (Marcus 2023; Wanless and Shapiro 2022).

- **IPCC for AI capabilities and safety assessment and policy recommendations:** The model similar to the Intergovernmental Panel on Climate Change (IPCC) should help provide scientific and technical guidance on AI technologies, their societal impacts and ways to mitigate its risks, particularly threats of biotechnologies and autonomous weapons (Ho et al. 2023). In addition, through this model, AI and

governance experts can work closely to provide fact-based advice on sustainable AI practices, assess AI systems regarding data and algorithm biases, and promote AI-driven solutions that are developed within ethical values and standard societal norms (Miailhe 2018). This model focuses more on AI's social and ethical implications and could help navigate complex dilemmas and unintended consequences of AI development. It could also bridge the gap between field experts and policy makers, ensuring AI is developed and used responsibly and ethically.

- **Geneva Convention-like protocol for AI:** Above all, considering past events of history in terms of violation of human rights and dignity, the newly formed global AI governance framework should consider Geneva Convention-like protocols based on international legal standards that develop and use AI to serve humanity and protect it from potential violations, conflicts and threats caused by AI. These threats could include deepfake issues, data abuse and misuse, algorithm bias, breach of individual and collective privacy, cybersecurity and surveillance (Horowitz 2018). Adopting an international treaty can help create specific ethical guidelines for different industries (for example, health care, finance, transportation) to ensure that AI is used in a way that aligns with the specific ethical considerations of that industry within a transparent, responsible and sustainable environment.

Implementing these recommendations presents a pathway for multilateral institutions to contribute actively to a responsible and ethical AI future. By embracing these roles, these institutions can ensure that AI technologies are developed and deployed in a manner that safeguards humanity and upholds international norms and values. Ultimately, the collaborative effort of these organizations will be pivotal in steering the global AI landscape toward sustainable, equitable and secure outcomes for all.

Conclusion

To comprehensively understand and shape the future trajectory of global governance of AI, a clear definition of it needs to be set, along with its societal impacts. The critical analysis of empirical and normative studies will shape the future global governance of AI and underscore the significance of multilateral cooperation in steering the direction of global AI governance. When AI is analyzed for how it changes societies, affects geopolitical domination struggles and pushes actors to visualize governance frameworks around it, it is possible to see how it will become a part of everyday life. By bringing states together through regulatory agreements involving authority, legitimacy and non-state actors, multilateral cooperation can mitigate too much competition within the AI governance space.

The future of multilateral AI governance should feature a harmonized framework of regulations and standards developed through collaborative international forums. This framework will prioritize ethical considerations, data privacy and human rights. It will foster innovation and stakeholder involvement, including states and civil society. If multilateral solutions do not work, the world will become more anarchic where states refuse to work together as realists would predict. It is critical to remain vigilant in ensuring that multilateral cooperation continues to be a practical approach even during impending crises involving AI supremacy, to develop a world where AI technology is used for human good.

Acknowledgements

About the Author

Maral Niazi is a Ph.D. student at the Balsillie School of International Affairs with a multidisciplinary background in political science, human rights, law, and global governance. Her research with the Digital Policy Hub will expand on her doctoral research on the global governance of AI where she will examine the societal impacts of AI on humanity.

Works Cited

Azoulay, Audrey. 2018. "Towards an Ethics of Artificial Intelligence." *UN Chronicle* 3 & 4 (LV). www.un.org/en/chronicle/article/towards-ethics-artificial-intelligence.

Bostrom, Nick and Eliezer Yudkowsky. 2014. "The Ethics of Artificial Intelligence." In *The Cambridge Handbook of Artificial Intelligence*, edited by Keith Frankish and William M. Ramsey, 316–34. Cambridge, UK: Cambridge University Press.

Butcher, James and Irakli Beridze. 2019. "What is the State of Artificial Intelligence Governance Globally?" *The RUSI Journal* 164 (5-6): 88–96. https://doi.org/10.1080/03071847.2019.1694260.

Carnegie Endowment for International Peace. 2019. *The Global Expansion of AI Surveillance*. Washington, DC: Carnegie Endowment for International Peace.

Cath, Corinne. 2018. "Governing artificial intelligence: ethical, legal and technical opportunities and challenges." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376: 20180080. https://doi.org/10.1098/rsta.2018.0080.

Cummings, Mary L., Heather M. Roff, Kenn Cukier, Hannah Bryce and Jacob Parakilas. 2018. *Artificial Intelligence and International Affairs: Disruption Anticipated*. Chatham House, June 14. https://apo.org.au/node/176326.

Donahoe, Eileen and Megan M. Metzger. 2019. "Artificial intelligence and human rights." *Journal of Democracy* 30 (2): 115–26. https://doi.org/10.1353/jod.2019.0029.

Erdélyi, Olivia J. and Judy Goldsmith. 2022. "Regulating artificial intelligence: Proposal for a global solution." *Government Information Quarterly* 39 (4): 101748. https://doi.org/10.1016/j.giq.2022.101748.

Erman, Eva and Markus Furendal. 2022. "The Global Governance of Artificial Intelligence: Some Normative Concerns." *Moral Philosophy and Politics* 9 (2): 267–91. https://doi.org/10.1515/mopp-2020-0046.

Floridi, Luciano. 2019. "What the Near Future of Artificial Intelligence Could Be." *Philosophy & Technology* 32 (1): 1–15. https://doi.org/10.1007/s13347-019-00345-y.

Floridi, Luciano and Josh Cowls. 2019. "A Unified Framework of Five Principles for AI in Society." SSRN, April 29. http://dx.doi.org/10.2139/ssrn.3831321.

Hamet, Pavel and Johanne Tremblay. 2017. "Artificial intelligence in medicine." *Metabolism, Clinical and Experimental* 69: 36–40. https://doi.org/10.1016/j.metabol.2017.01.011.

Hendrix, Justin. 2023. "Exploring Global Governance of Artificial Intelligence." Tech Policy Press, June 25. https://techpolicy.press/exploring-global-governance-of-artificial-intelligence/.

Ho, Lewis, Joslyn Barnhart, Robert Trager, Yoshua Bengio, Miles Brundage, Allison Carnegie, Rumman Chowdhury et al. 2023. "International Institutions for Advanced AI." ArXiv. https://doi.org/10.48550/arXiv.2307.04699.

Hoffmann-Riem, Wolfgang. 2019. "Artificial Intelligence as a Challenge for Law and Regulation." In *Regulating Artificial Intelligence*, edited by Thomas Wischmeyer and Timo Rademacher, 1–29. Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-030-32361-5_1.

Horowitz, Michael C. 2018. "Artificial Intelligence, International Competition, and the Balance of Power." *Texas National Security Review* 1 (3): 36–57. https://doi.org/10.15781/T2639KP49.

Hosny, Ahmed, Chintan Parmar, John Quackenbush, Lawrence H. Schwartz and Hugo J. Aerts. 2018. "Artificial intelligence in radiology." *Nature Reviews Cancer* 18 (8): 500–10. https://doi.org/10.1038/s41568-018-0016-5.

IBM. 2018. "AI Fairness 360." IBM Developer. November 14. www.ibm.com/opensource/open/projects/ai-fairness-360/.

Janiesch, Christian, Patrick Zschech and Kai Heinrich. 2021. "Machine learning and deep learning." *Electronic Markets* 3 (31): 685–95. https://doi.org/10.1007/s12525-021-00475-2.

Jobin, Anna, Marcello Ienca and Effy Vayena. 2019. "The global landscape of AI ethics guidelines." *Nature Machine Intelligence* 1 (9): 389–99. https://doi.org/10.1038/s42256-019-0088-2.

Liebig, Laura, Licinia Güttel, Anna Jobin and Christian Katzenbach. 2022. "Subnational AI policy: shaping AI in a multi-level governance system." *AI & Society*. https://doi.org/10.1007/s00146-022-01561-5.

Lopez-Claros, Augusto, Arthur L. Dahl and Maja Groff. 2020. *Global Governance and the Emergence of Global Institutions for the 21st Century*. Cambridge, UK: Cambridge University Press.

Marcus, Gary. 2023. "A CERN for AI and the Global Governance of AI." *Marcus on AI*, October 26. https://garymarcus.substack.com/p/a-cern-for-ai-and-the-global-governance.

Miailhe, Nicolas. 2018. "Why We Need an Intergovernmental Panel for Artificial Intelligence." Our World, United Nations University, December 21. https://ourworld.unu.edu/en/why-we-need-an-intergovernmental-panel-for-artificial-intelligence.

Miailhe, Nicholas and Yolanda Lannquist. 2020. "Global Governance of Artificial Intelligence." In *Handbook of Artificial Intelligence and Robotic Process Automation: Policy and Government Applications*, edited by Al Naqvi and J. M. Munoz. London, UK: Anthem Press.

Mäntymäki, Matti, Matti Minkkinen, Teemu Birkstedt and Mika Viljanen. 2022. "Defining organizational AI governance." *AI and Ethics* 2 (4): 603–9. https://doi.org/10.1007/s43681-022-00143-x.

Nichols, Michelle. 2023. "UN chief backs idea of global AI watchdog like nuclear agency." Reuters, June 12. www.reuters.com/technology/un-chief-backs-idea-global-ai-watchdog-like-nuclear-agency-2023-06-12/.

NSTC. 2023. *The National Artificial Intelligence Research and Development Strategic Plan*. Select Committee on Artificial Intelligence Report. The Office of the President of the United States. May. www.whitehouse.gov/wp-content/uploads/2023/05/National-Artificial-Intelligence-Research-and-Development-Strategic-Plan-2023-Update.pdf.

OECD. 2019. "Recommendation of the Council on Artificial Intelligence." OECD Legal Instruments. https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449.

———. 2023. "Human-centered Values and Fairness." OECD.AI Policy Observatory. https://oecd.ai/en/dashboards/ai-principles/P6.

PAI. 2020. "PAI Researchers Co-author Multistakeholder Report on Improving Verifiability in AI Development." *Partnership on AI* (blog), November 3. https://partnershiponai.org/pai-researchers-co-author-multistakeholder-report-on-verifying-claims-for-ai-development/.

P. S., Varsha. 2023. "How can we manage biases in artificial intelligence systems – A systematic literature review." *International Journal of Information Management Data Insights* 3 (1): 100165. https://doi.org/10.1016/j.jjimei.2023.100165.

Rhodes, R. A. 1996. "The New Governance: Governing without Government." *Political Studies* 44 (4): 652–67. https://doi.org/10.1111/j.1467-9248.1996.tb01747.x.

Rosenau, James N. and Ernst-Otto Czempiel. 1992. *Governance Without Government: Order and Change in World Politics*. Cambridge, UK: Cambridge University Press.

Russell, Stuart and Peter Norvig. 2016. *Artificial Intelligence: A Modern Approach*, 3rd ed. London, UK: Pearson Education.

Schmitt, Lewin. 2021. "Mapping global AI governance: a nascent regime in a fragmented landscape." *AI and Ethics* 2 (2): 303–14. https://doi.org/10.1007/s43681-021-00083-y.

Sepasspour, Rumtin. 2023. "A reality check and a way forward for the global governance of artificial intelligence." *Bulletin of the Atomic Scientists* 79 (5): 304–15. https://doi.org/10.1080/00963402.2023.2245249.

Tallberg, Jonas, Eva Erman, Markus Furendal, Johannes Geith, Mark Klamberg and Magnus Lundgren. 2023. "The Global Governance of Artificial Intelligence: Next Steps for Empirical and Normative Research." *International Studies Review* 25 (3): viad040, https://doi.org/10.1093/isr/viad040.

Ulnicane, Inga, William Knight, Tonii Leach, Bernd C. Stahl and Winter-Gladys Wanjiku. 2022. "Governance of Artificial Intelligence: Emerging International Trends and Policy Frames." In *The Global Politics of Artificial Intelligence*, edited by Maurizio Tinnirello, 29–55. Boca Raton, FL: CRC Press.

UN Framework Convention on Climate Change. 2023. "AI for Climate Action: Technology Mechanism supports transformational climate solutions." November 3. https://unfccc.int/news/ai-for-climate-action-technology-mechanism-supports-transformational-climate-solutions.

UNESCO. 2019. "Preliminary study on a possible standard-setting instrument on the ethics of artificial intelligence." Paper presented at UNESCO 40th General Conference, Paris, France.

Vinuesa, Ricardo, Hossein Azizpour, Iolanda Leite, Madeline Balaam, Virginia Dignum, Sami Domisch, Anna Felländer, Simone D. Langhans, Max Tegmark and Francesco Fuso Nerini. 2020. "The role of artificial intelligence in achieving the Sustainable Development Goals." *Nature Communications* 11 (1): 233. https://doi.org/10.1038/s41467-019-14108-y.

Voigt, Paul and Axel von dem Bussche. 2017. *The EU General Data Protection Regulation (GDPR): A Practical Guide*. Basingstoke, UK: Springer.

Wanless, Alicia and Jacob N. Shapiro. 2022. "A CERN Model for Studying the Information Environment." Carnegie Endowment for International Peace. https://carnegieendowment.org/2022/11/17/cern-model-for-studying-information-environment-pub-88408.

Webster, Graham, Rogier Greemers, Elsa Kania and Paul Triolo. 2017. "Full Translation: China's 'New Generation Artificial Intelligence Development Plan' (2017)." DigiChina, Stanford University. https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/.

Weiss, Thomas G. 2016. *Global Governance: Why? What? Whither?* Hoboken, NJ: John Wiley & Sons.

Zeng, Jinghan. 2022. "China's Authoritarian Governance and AI." *International Affairs* 96 (6): 1441–59. https://doi.org/10.1093/ia/iiaa172.

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York, NY: Public Affairs.