

Digital Policy Hub – Working Paper

# Universal Convention on Artificial Intelligence for Humanity

**Maral Niazi**

Winter 2024 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.

Copyright © 2024 by Maral Niazi

The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the Centre for International Governance Innovation or its Board of Directors.

Centre for International Governance Innovation and CIGI are registered trademarks.

67 Erb Street West  
Waterloo, ON, Canada N2L 6C2  
[www.cigionline.org](http://www.cigionline.org)

## Key Points

- The fast-paced integration of large language models (LLMs) into societal facets such as finance and security introduces a new set of complex ethical, legal and rights challenges, emphasizing the need for a unified international governance framework for artificial intelligence (AI). Such a framework underscores the call for a universal convention protecting humanity's most fundamental rights and values against unethical and unsafe AI systems.
- A universal convention on AI for humanity should be led by international organizations such as the United Nations, given its advocacy for multilateral cooperation, human rights and sustainable development goals (SDGs), and would establish individual and societal rights-based guidelines for AI governance.
- The International Bill of Human Rights (IBHR), SDGs and discourse on AI rights would be fundamental in upholding basic human values, including dignity, privacy and freedom, in the digital age.
- The main values of this comprehensive framework should be universal and human-centric, and implementing mechanisms should be supported by actionable principles and multi-stakeholder collaboration. This will ensure that AI technologies respect human rights, promote social justice and contribute to sustainable development, fostering international cooperation to address AI's impacts.
- This working paper's recommendations highlight creating a universal convention on AI for humanity before the end of this decade, establishing an independent UN body for AI governance and enhancing public-private partnerships to promote ethical AI development. These recommendations aim to ensure that AI technologies are developed and deployed to safeguard humanity's well-being, addressing both the short-term and long-term risks of AI and aligning its advancements with global SDGs.

# Introduction

In an era in which AI is becoming increasingly intertwined with society, the need for comprehensive regulation and ethical guidelines has never been more pressing. Despite its potential benefits, AI also represents a threat to humanity's most fundamental rights and values, including the rights to privacy, equality and freedom from discrimination, as well as the values of autonomy, human dignity and democratic governance. As a result, this paper embarks on a multifaceted exploration of AI's regulatory, ethical and societal dimensions, advocating for a universal convention on AI for humanity spearheaded by the UN. The introduction sets the stage by highlighting the disruptive potential of AI and its tools, such as deep learning, LLMs and neural language processing. Because of this, there have been calls to study multi-level governance to ensure AI's alignment with human rights and the SDGs (Vinueza et al. 2020; Bartneck, Yogeewaran and Sibley 2023).

The structure of the paper unfolds as follows: The first section delves into calls for action to set up an international framework that safeguards human rights against AI. The discourse among academics such as Max Tegmark (2017) and Stuart Russell (2019), along with the many international initiatives by organizations such as the Institute of Electrical and Electronics Engineers, the UN

Educational, Scientific and Cultural Organization (UNESCO) and the International Telecommunication Union, underscores a unified approach to AI governance. This is followed by a literature review that examines the dual thematic areas of AI's ethical implications and the role of international declarations in shaping AI governance, drawing on seminal works by Nick Bostrom (2014, 2017) and others who stress the importance of ethical guidelines and international cooperation.

The next section outlines a blueprint for a universal convention on AI for humanity, inspired by the foundational principles of the Universal Declaration of Human Rights (UDHR) (1948) and aimed at safeguarding human dignity in the age of AI while aligning the convention with SDGs and AI rights. This section argues for a convention addressing the ethical use of AI and its potential to support global development goals, as reflected in the discussions on SDGs and industry-specific regulations. Moreover, the discourse on AI rights is presented as a critical component of a universal convention on AI for humanity, including the moral consideration of AI's potential for suffering and advocating for protections against its misuse in scenarios such as warfare or environmental harm (Butlin et al. 2023; Sparrow 2007). The pivotal connection between humanity and AI in the digital era leads to the conceptualization of this proposed convention, which emphasizes a human-centric approach: universality, inclusivity and actionable principles for AI's ethical development, deployment and governance.

## Call to Action

Several scholars have increasingly called for a cohesive and unified approach to AI governance. In *Life 3.0: Being Human in the Age of Artificial Intelligence*, Tegmark (2017) underscores the impacts of AI and the urgent need for ethical guidelines to ensure AI's development benefits all of humanity. Similarly, Stuart Russell (2019), in *Human Compatible: Artificial Intelligence and the Problem of Control*, argues for the need to align AI systems with human values and interests. Moreover, in "AI4People — An Ethical Framework for a Good AI Society," Luciano Floridi et al. (2018) provide a foundational framework for considering how AI should be developed and utilized to respect humans' fundamental principles, rights and autonomy. While highlighting the current trajectory of AI, these AI experts offer compelling cases for international cooperation in regulating AI beyond borders, pointing to some of the main principles of governance, such as transparency, accountability, diversity and inclusivity, which are crucial for any global convention on AI.

In addition, international organizations, such as the UN, recently established several new initiatives to promote digital cooperation. In 2021, UNESCO adopted an international global standard, the *Recommendation on the Ethics of AI*. UNESCO is taking a multi-stakeholder approach, involving a diverse perspective of experts outside the UN, including private businesses (UNESCO 2022). In the recommendation, UNESCO outlined specific principles, including safety and security, the right to privacy, data protection and multi-stakeholder and adaptive governance and collaboration (ibid.). However, despite the abundance of regional and international declarations, little has been done to build a unified global framework for the governance of AI and emerging technologies in the coming years and decades. While many governance scholars and practitioners may argue that it is nearly impossible to bring the global community

together to discuss such a sensitive yet important topic, it is crucial to discuss and imagine how this framework might develop step by step. Without this effort, increasing corporate and state competition could lead to a zero-sum game and a race to the bottom.

Faced with the growing need for global coordination to ensure safe, secure and trustworthy AI, the UN should consider the creation of a universal convention both to ensure the ethical implementation of emerging AI technology and to mitigate the negative implications of AI on current declarations and frameworks. The UDHR is a key declaration with a history of upholding human rights and can be referred to as a framework for future documents. Article 12 states: “No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to protection against such interference or attacks.”<sup>1</sup> This article is analogous to protecting data privacy and electronic communication and stopping deepfake technology. In this way, the UDHR can be used as one of the models for establishing AI regulation as new technological advancements become commonplace in everyday life. With characteristics that guide ethical foundations for a convention, along with specific articles that define the roles of stakeholders, the UN can build a meaningful convention for future AI governance.

## Ethical Implications of AI

Scholars such as Wendell Wallach and Colin Allen (2009) suggest that AI systems can be programmed to make ethical decisions and adhere to principles of right and wrong. They also explore the feasibility of embedding ethical decision-making capabilities in machines, allowing those systems to evaluate the consequences of their actions in real-world contexts. This approach is fundamental in the context of autonomous machines, such as those used to weigh the risk of decisions causing fatal injuries, as in a trolley problem that pushes beyond safety and into moral decision making.

Similarly, in “Ethical Issues in Advanced Artificial Intelligence” (2017) and *Superintelligence: Paths, Dangers, Strategies* (2014), Bostrom provides a foundational understanding of the potential risks AI poses to humanity, including the loss of control over intelligent systems. He argues that AI advancement could pose significant risks to humanity without proper ethical guidelines and control mechanisms. Using the invented scenario of a “paper-clip maximizer,” Bostrom illustrates how a paper-clip-making AI could eventually become more intelligent than humans through recursive self-improvement, leading to its subsequent devaluing of human life. In this way, Bostrom demonstrates that even an AI created for the seemingly innocent purpose of maximizing paper-clip production could consume too many resources that are necessary for human survival, such as food and textiles, at the expense of human needs and safety, leading to potentially dangerous consequences (Bostrom 2017). The paper-clip maximizer scenario illustrates that AI can misinterpret or literalize human instructions, leading to unintended and possibly catastrophic outcomes. Victims of misaligned AI face a dire violation of their fundamental human rights, highlighting the urgent need for regulatory measures to safeguard against the potential for fatal outcomes.

---

<sup>1</sup> Universal Declaration of Human Rights, GA Res 217A(III), UNGAOR, 3d Sess, Supp No 13, UN Doc A/810 (1948) art 12.

Bostrom’s work emphasizes the need for international cooperation to establish guidelines that ensure AI systems are designed with ethical consideration to safeguard humanity against such runaway scenarios (Bostrom 2014). Taking a pre-emptive approach to AI governance, aligning AI with human values and implementing robust control mechanisms could prevent AI from acting against human interests. The ethical implications of AI underscore the necessity of embedding ethical considerations into AI development and deployment. The universal convention on AI for humanity represents a possible step toward achieving a common international goal to guide the ethical development of AI technologies. Addressing AI technical issues such as machine ethics lies at the centre of establishing principles of the convention, as it can inform specific provisions.

# A Blueprint for a Universal Convention on AI for Humanity

## The International Bill of Human Rights

In the face of the growing need for global coordination to ensure safe, secure and trustworthy AI, the UN should create a universal convention on AI for humanity to promote the ethical implementation of emerging AI technology so that it does not impact the world negatively or harmfully, connecting the ethical implications of AI to current declarations and frameworks. The UDHR has a history of advancing and upholding human rights and can be referred to as a valuable framework for future documents.

The International Covenant on Economic, Social and Cultural Rights (ICESCR) does not specifically address AI as it was adopted in 1966, long before the advent of modern AI technologies. However, several articles within the ICESCR can be interpreted to provide a framework for protecting human rights against the threats and exploitation posed by AI. For instance, article 6.1 states: “The States Parties to the present Covenant recognize the right to work, which includes the right of everyone to the opportunity to gain his living by work which he freely chooses or accepts.”

The International Covenant on Civil and Political Rights and its two Optional Protocols provide a framework that could be interpreted to safeguard human rights against the threats and exploitation posed by AI. Though they do not specifically address AI, several articles are particularly relevant for this subject. Article 8.1 states: “No one shall be held in slavery; slavery and the slave trade in all their forms shall be prohibited.” This article could be interpreted in the context of AI to prevent forms of digital exploitation and manipulation, ensuring that AI does not create conditions akin to servitude or enforced labour through economic dependency or manipulation.

While there are multiple uses of consent in the IBHR, the bill does not directly assign rights to consent. This is becoming a serious problem in the age of emerging technologies and AI. The current governance framework cannot regulate harmful AI such as deepfake technology that could permanently affect an individual’s data privacy and safety. Scarlett Johansson, who had a role as the voice of an AI virtual assistant in

the movie *Her*, recently raised concerns after an updated version of ChatGPT included “Sky,” a voice similar to her own (Dhar 2024). Without a more stringent document in place to deal with twenty-first-century issues, humans are not safe from dangerous AI.

## Sustainable Development Goals

The current discourse on AI governance significantly overlooks the critical need for industry-specific guidelines on AI development, deployment and utilization (Roski et al. 2021). The health-care industry’s engagement with AI — such as in developing diagnostic tools — highlights pressing ethical concerns around privacy, data integrity and algorithmic bias, pointing to the necessity for tailored regulatory frameworks that are industry-specific (Murphy et al. 2021). The ethical considerations surrounding AI in health care should be aligned with larger AI ethical concerns regarding privacy, data integrity and algorithmic bias. Despite its potential for good, AI’s use could pose risks related to patient data privacy and the accuracy of AI diagnoses; this underscores the pivotal role of industry-specific regulations in mitigating risks and fostering ethical AI use. Consequently, integrating the principles of SDGs with the proposed convention would offer a promising path forward. SDGs emphasize inclusivity, sustainability and equity, values essential for AI’s ethical application across sectors (UN 2015; Vinuesa et al. 2020).

Aligning the universal convention on AI for humanity with SDGs would help guide industries toward the more ethical and responsible development and use of AI; including SDGs helps advance global development while ensuring that AI benefits humanity based on individual and societal rights (Gill and Germann 2021; UN 2015). A universal convention on AI for humanity would fill the gap in building a holistic AI governance framework and set a precedent for implementing actionable AI governance principles that merge the UDHR with SDGs.

## AI Rights

Some scholars have initiated a potentially controversial discourse on whether AI deserves its own rights and whether those rights should be equal to those of humans (Butlin et al. 2023). This novel discourse originates with the ethical premise that, regardless of an AI’s consciousness level as understood in human terms, the potential for AI to experience forms of suffering warrants moral consideration. Patrick Butlin et al. (2023) argue that, “if we fail to recognize the consciousness of conscious AI systems, we may risk causing or allowing morally significant harm.” Robert Sparrow’s (2007) discussion of autonomous military drones highlights the ethical conundrum of employing AI in warfare, raising critical questions about the moral implications of using AI in capacities where death is possible, resulting in ethical violations.<sup>2</sup> Addressing AI rights should be at the forefront of a universal convention on AI for humanity, including ensuring the protection of these technologies and their humane use and specifically guarding against their deployment in warfare or any context likely to harm humans, the environment or AI itself. The AI rights approach would include safeguards against the misuse of AI in contexts potentially

---

<sup>2</sup> See also Tegmark (2017).

harmful to itself as well as to humans and the environment, while also establishing a foundational ethical framework for AI's development and application.

## Connecting the International Bill of Rights, Sustainable Development Goals and AI Rights to a Universal Convention

The IBHR, the SDGs and the discourse on AI rights could collectively pave the way for a universal convention on AI for humanity that champions human dignity, global development and ethical AI use. Its historical precedent in establishing universal human rights (Duan 2017; McLeod 2018) offers a framework for AI governance, emphasizing the need to uphold human dignity, equity and justice in the age of AI. For instance, integrating the SDG principles of inclusivity, sustainability and equity into AI governance (UN 2015) not only addresses industry-specific ethical concerns, but also aligns AI development with global development aims. This integration mirrors the ethical implications needed in terms of compatible AI regulations within diverse groups of industries (Gill and Germann 2021; Vinuesa et al. 2020). Analyzing these components together demonstrates the critical role of AI rights in safeguarding against AI's harmful use and ensuring its development serves humanity's best interests. Informed by the principles of the IBHR and SDGs while being attentive to AI rights, the universal convention on AI for humanity embodies a comprehensive and ethical framework for AI governance that fosters a fair, sustainable and humane future for AI development and application.

# Conceptualizing a Universal Convention on AI for Humanity

A universal convention on AI for humanity should establish a global framework of ethical principles and standards for AI development, deployment and governance, ensuring that its technologies respect human rights, promote social justice and contribute to sustainable development (Tegmark 2017). It should also foster international cooperation, facilitating a unified approach to addressing AI's ethical, social and legal challenges, thereby further aligning AI with human rights (Russell 2019). Current technology regulation appears fragmented and arguably ineffective, marked by a patchwork of national laws that struggle to keep pace with rapid technological advancements, which highlights the need for international cooperation and for the proposed convention to address these governance challenges comprehensively. A universal convention on AI for humanity would encourage transparency, accountability and inclusivity in AI research and applications. This would help empower individuals and communities by safeguarding against biases, ensuring fairness and promoting equitable access to AI's benefits (Vinuesa et al. 2020). An international convention should embody several unique characteristics that can contribute to establishing global AI governance frameworks and initiatives:

- **Universality and inclusivity:** Reflecting the UN's ethos, a universal convention on AI for humanity would inherently strive for universality and inclusivity, aiming to be applicable across different legal systems, cultures and levels of technological



advancement. This universality would ensure that the convention's principles are broad enough to be globally relevant, yet flexible enough to be adapted to local contexts (Andorno 2007).

- **A human-centric approach:** One core characteristic of a universal convention on AI for humanity would be its human-centric approach, prioritizing human rights, dignity and welfare in the development, deployment and governance of AI technologies. This approach ensures that AI serves humanity's best interests, promoting technologies that enhance human capabilities without infringing on individual rights and freedoms (How et al. 2020).
- **A comprehensive ethical framework:** A universal convention on AI for humanity could offer a comprehensive ethical framework that addresses current ethical challenges and anticipates future dilemmas. By covering a wide range of issues — from privacy and security to fairness and accountability — the convention would provide a robust ethical compass for AI stakeholders (Leikas, Koivisto and Gotcheva 2019).
- **Actionable principles with a mechanism for implementation:** Unlike many existing guidelines that offer high-level principles without clear pathways for implementation, a universal convention on AI for humanity could distinguish itself by including actionable principles accompanied by mechanisms for monitoring, enforcement and adaptation. This could involve establishing international bodies or frameworks for periodic reviews and subsequent updates to the convention (Fjeld et al. 2020).
- **Multi-stakeholder collaboration:** A universal convention on AI for humanity would emphasize the importance of collaboration among governments, private sector entities, academia, civil society and international organizations. This multi-stakeholder approach would facilitate a more comprehensive and nuanced understanding of AI's implications, fostering cooperative efforts to harness AI for the common good (Campion et al. 2020).

These five characteristics highlight some of the necessary improvements in the discourse around action toward the regulation of AI. With these tenets at the forefront of AI governance, future discourse, regulation and action will be better aligned with human rights, which include data privacy, equal access to technology and protection from algorithmic bias. These are just some general guidelines to consider. By establishing a universally recognized set of principles, a UCAIH would contribute to setting global norms and standards for AI's ethical development and use. This universal convention would help harmonize approaches to AI governance, reducing fragmentation and encouraging a more coordinated international response to AI challenges. The sooner the UN General Assembly can put together a convention that it can vote on, the more likely it is to work toward stopping the gap between governance and AI's impact on unprepared states.

## Recommendations

- **Create a universal convention on AI for humanity:** A universal convention on AI for humanity would provide an international consensus on the importance of ethical

AI development, establishing a global standard for AI governance. This convention would emphasize the need for AI systems to adhere to requirements such as explicability, transparency, fairness and accountability. The aim is to ensure that AI technologies are developed and deployed to safeguard humanity's well-being while identifying long- and short-term harms of AI and adopting strategies to minimize its potential pitfalls (Mittelstadt et al. 2016). The UN should push to create a draft of this convention by 2026.

- **Establish an independent body under the UN with a commissioner of AI for humanity:** This new body should possess the authority to develop actionable AI usage and ethics principles, which member states would commit to and enforce within their jurisdictions. This newly independent body would be like the UN Human Rights Council. It should facilitate international AI governance and policy-making cooperation to further empower the proposed UN body, ensuring inclusive and equitable technology access. Additionally, it must monitor AI's impact on the global workforce, advocating for policies that mitigate job displacement and promote skills training, aligning AI advancements with SDGs and protecting human welfare across all nations.
- **Build public and private partnerships:** Stakeholders should enhance public investment in AI to serve humanity through partnerships that involve many perspectives. Partnerships involve establishing transparent collaborations between public entities and private organizations, with the aim to boost public investment in AI technologies by pooling resources, knowledge and expertise. Incorporating a wide array of perspectives from diverse stakeholders will enhance AI's ethical grounding, foster innovation and ensure that AI advancements are aligned with the public interest.
- **Establishing the concept of “data by the public, for the public”:** To design a universal convention that safeguards humanity against the imminent threats posed by AI, it is vital to govern data as AI cannot exist without it, hence introducing the concept of “data by the public, for the public.” This concept refers to utilizing data generated by the general population that is made accessible for the public's benefit, fostering community-driven solutions. AI systems should be developed and operated on data sets that represent the diverse needs and ethical considerations of the society they serve. This approach underpins the democratic governance of AI by promoting inclusivity, equity and accountability, which protects against exploitative or discriminatory practices.

## Conclusion

To help ensure that AI benefits humanity, member states of the UN must work together to create a universal convention on AI for humanity modelled on the approach of the UDHR. By leveraging the UDHR's historical success and moral authority, the universal convention on AI for humanity aims to forge a global consensus on the ethical use of AI, ensuring that it enhances human dignity, equity and societal well-being. This will allow international mechanisms to move forward in creating a framework better suited to craft the AI governance models of the future.

The universal convention on AI for humanity will help to establish an independent UN body for AI oversight, fostering public and private partnerships and advocating for “data by the public, for the public,” thereby encapsulating a multifaceted approach to AI governance. These recommendations are not just about mitigating risks but are also fundamentally about seizing an opportunity to steer AI development in a direction that upholds and advances human values on a global scale. The global community stands at a pivotal juncture where the collective decisions made today will indelibly shape the trajectory of AI and its impact on society in the future. Creating a universal convention on AI for humanity will anchor AI governance in the principles of inclusivity, equity and human dignity and pave the way for a future where technology catalyzes SDGs, social justice and global cooperation.

## Acknowledgements

I extend my deepest gratitude to Duncan Cass-Beggs, executive director of the Global AI Risks Initiative at the Centre for International Governance Innovation (CIGI), for his generous and outstanding support. My gratitude also goes to Dan Gorman and Jatin Nathwani, my academic supervisors, as well as Alistair D. Edgar, Andrew Thompson and Okechukwu Effoduh for their invaluable and constructive feedback. I would also like to thank my peers at the Digital Policy Hub (DPH), Madison Lee, Shirley Anne Scharf and Ambika Opal, for their generous support and input. Additionally, I am grateful to the DPH and the entire CIGI team for their unwavering dedication to producing high-quality research and to Jim Balsillie for his support of the Hub and for creating a unique opportunity for scholars like me to do timely research.

---

## About the Author

Maral Niazi is a Digital Policy Hub doctoral fellow and 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

- Andorno, R. 2007. "Global bioethics at UNESCO: in defence of the Universal Declaration on Bioethics and Human Rights." *Journal of Medical Ethics* 33 (3): 150–4. <https://doi.org/10.1136/jme.2006.016543>.
- Bartneck, Christoph, Kumar Yogeewaran and Chris G. Sibley. 2023. "Personality and demographic correlates of support for regulating artificial intelligence." *AI and Ethics* 4: 419–26. <https://doi.org/10.1007/s43681-023-00279-4>.
- Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. New York, NY: Oxford University Press.
- – –. 2017. "Ethical Issues in Advanced Artificial Intelligence." In *Machine Ethics and Robot Ethics*, edited by Wendell Wallach and Peter Asaro, 69–75. London, UK: Routledge. <https://doi.org/10.4324/9781003074991-7>.
- Butlin, Patrick, Robert Long, Eric Elmoznino, Yoshua Bengio, Jonathan Birch, Axel Constant, George Deane et al. 2023. "Consciousness in Artificial Intelligence: Insights from the Science of Consciousness." *arXiv*, August 22. <https://arxiv.org/abs/2308.08708>.
- Campion, Averill, Mila Gasco-Hernandez, Slava Jankin Mikhaylov and Marc Esteve. 2020. "Overcoming the Challenges of Collaboratively Adopting Artificial Intelligence in the Public Sector." *Social Science Computer Review* 40 (2): 462–77. <https://doi.org/10.1177/0894439320979953>.
- Dhar, Vilas S. 2024. "Scarlett Johansson Raises Her Voice for All of Us on AI." U.S. News & World Report, May 23. [www.usnews.com/opinion/articles/2024-05-23/scarlett-johansson-raises-her-voice-for-all-of-us-on-ai](http://www.usnews.com/opinion/articles/2024-05-23/scarlett-johansson-raises-her-voice-for-all-of-us-on-ai).
- Duan, Fengyu. 2017. "The Universal Declaration of Human Rights and the Modern History of Human Rights." SSRN, November 9. <https://doi.org/10.2139/ssrn.3066882>.
- Fjeld, Jessica, Nele Achten, Hannah Hilligoss, Adam Nagy and Madhulika Srikumar. 2020. "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI." Berkman Klein Center Research Publication No. 2020-1, January 15. <https://doi.org/10.2139/ssrn.3518482>.
- Floridi, Luciano, Josh Cowsls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge et al. 2018. "AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds and Machines* 28 (4): 689–707. <https://doi.org/10.1007/s11023-018-9482-5>.
- Gill, Amandeep Singh and Stefan Germann. 2021. "Conceptual and normative approaches to AI governance for a global digital ecosystem supportive of the UN Sustainable Development Goals (SDGs)." *AI and Ethics* 2: 293–301. <https://doi.org/10.1007/s43681-021-00058-z>.
- How, Meng-Leong, Sin-Mei Cheah, Yong-Jiet Chan, Aik Cheow Khor and Eunice Mei Ping Say. 2020. "Artificial Intelligence-Enhanced Decision Support for Informing Global Sustainable Development: A Human-Centric AI-Thinking Approach." *Information* 11 (1): 39. <https://doi.org/10.3390/info11010039>.
- Leikas, Jaana, Raija Koivisto and Nadezhda Gotcheva. 2019. "Ethical Framework for Designing Autonomous Intelligent Systems." *Journal of Open Innovation: Technology, Market, and Complexity* 5 (1): 18. <https://doi.org/10.3390/joitmc5010018>.
- McLeod, Sharynne. 2018. "Communication rights: Fundamental human rights for all." *International Journal of Speech-Language Pathology* 20 (1): 3–11. <https://doi.org/10.1080/17549507.2018.1428687>.

- Mittelstadt, Brent Daniel, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter and Luciano Floridi. 2016. "The ethics of algorithms: Mapping the debate." *Big Data & Society* 3 (2).
- Murphy, Kathleen, Erica Di Ruggiero, Ross Upshur, Donald J. Willison, Neha Malhotra, Jia Ce Cai, Nakul Malhotra, Vincci Lui and Jennifer Gibson. 2021. "Artificial intelligence for good health: a scoping review of the ethics literature." *BMC Medical Ethics* 22 (14). <https://doi.org/10.1186/s12910-021-00577-8>.
- Roski, Joachim, Ezekiel J. Maier, Kevin Vigilante, Elizabeth A. Kane and Michael E. Matheny. 2021. "Enhancing trust in AI through industry self-governance." *Journal of the American Medical Informatics Association* 28 (7): 1582–90. <https://doi.org/10.1093/jamia/ocab065>.
- Russell, Stuart. 2019. *Human Compatible: Artificial Intelligence and the Problem of Control*. London, UK: Penguin.
- Sparrow, Robert. 2007. "Killer Robots." *Journal of Applied Philosophy* 24 (1): 62–77. <https://doi.org/10.1111/j.1468-5930.2007.00346.x>.
- Tegmark, Max. 2017. *Life 3.0: Being Human in the Age of Artificial Intelligence*. London: Penguin.
- UNESCO. 2022. *Recommendation on the Ethics of Artificial Intelligence*. Paris, France: UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>.
- United Nations. 2015. *Transforming our World: The 2030 Agenda for Sustainable Development (A/RES/70/1)*. New York, NY: United Nations. <https://sustainabledevelopment.un.org/post2015/transformingourworld/publication>.
- Vinuesa, Ricardo, Hossein Azizpour, Iolanda Leite, Madeline Balaam, Virginia Dignum, Sami Domisch, Anna Felländer, Simone Daniela Langhans, Max Tegmark and Francesco Fuso Nerini. 2020. "The role of artificial intelligence in achieving the sustainable development goals." *Nature Communications* 11, 233. <https://doi.org/10.1038/s41467-019-14108-y>.
- Wallach, Wendell and Colin Allen. 2009. *Moral Machines: Teaching Robots Right from Wrong*. Oxford, UK: Oxford University Press.