

**HUMAN RIGHTS
AND
DIGITAL
TECHNOLOGY**

Submission to the
U.N. High-Level Advisory Body on Artificial Intelligence

in response to the
Call for Papers on Global AI Governance

28 September 2023

Contributors:

Professor Susan PERRY, a specialist in international human rights law, directs the graduate program in Diplomacy and International Law at The American University of Paris. A member of the French National Human Rights Commission (CNCDH), she has served on the advisory board of a recent European Union project on ethical AI. Contact: sperry@aup.edu

Professor Claudia RODA is professor of Computer Science, Director of the Human Rights and Data Science program, and former Dean at The American University of Paris. Her research focuses on the impact of digital technology on human behavior and social structures. She is the co-founder, with Professor Susan Perry, of the Research Center on Human Rights and Digital Technology. Contact: croda@aup.edu

Nicole SANTIAGO, an international human rights lawyer, is a Research Manager for law and emerging technologies at Trilateral Research and research partner at the Research Center on Human Rights and Digital Technology. Contact: nicoleelizabethsantiago@gmail.com

The *Research Center on Human Rights and Digital Technology* welcomes the opportunity to provide input to this call on global AI Governance. Based on our expertise on the impacts of digital technology on human rights, we offer the following recommendations:

Global AI Governance must be framed within international human rights.

The protection of human rights must be the upmost priority for global AI governance. Human rights are inherent to all human beings. AI, as a tool developed by humans, must not be used in ways that violate or interfere with any human being's inherent human rights.

Furthermore, there are practical reasons that human rights should be the governing framework, as opposed to other frameworks such as 'ethics' or 'trustworthiness'. Human rights is not a vague concept; it is an established framework with legally binding obligations. All U.N. Member states have human rights obligations stemming from the U.N. Charter and international human rights treaties, as well as

regional and national law. Additionally, human rights are enforceable through mechanisms at the international, regional, and national level. And furthermore, human rights are not static. The framework continuously evolves to respond to contemporary challenges. Thanks to the interpretative work of designated entities like treaty monitoring committees and the Special Procedures of the U.N. Office of the High Commissioner for Human Rights (OHCHR), we have a growing body of interpretative guidance that helps Member States understand their human rights obligations.

That being said, existing human rights monitoring and enforcements mechanisms are not resourced enough to deal with the wide-spread concerns posed by AI. Dedicated resources, training and fundings are urgently needed, including for the OHCHR.

Global AI Governance should consider impacts throughout AI hardware supply chains.

AI has many over-looked adverse environmental impacts. These will only continue to multiply as AI-uptake is encouraged and AI becomes used more broadly. Such adverse impacts include the use of toxic and non-biodegradable components in AI-enabled devices and high energy consumption. Consideration of these adverse aspects much span the entire lifespan of AI hardware, from mining the raw materials to manufacturing to disposal and recycling.

Global AI Governance should consider impacts throughout entire lifecycle of AI software.

AI governance should play a crucial role in every stage of the life cycle of AI software, from development and deployment to ongoing monitoring and maintenance. It must establish policies, procedures, and guidelines to ensure that AI systems respect human rights. The development and training processes should ensure that data collection and preparation are based on representative and unbiased datasets, adhere to privacy regulations, and are properly documented to safeguard transparency. Guidelines should be in place to guide the selection of algorithms, models, and techniques that align with ethical and legal considerations. The testing and validation processes must ensure the AI system performs as intended, not only in terms of accuracy, precision, recall, and other relevant metrics, but also in the protection of human rights. The deployment process must ensure compliance with regulations, in particular with policy requiring that AI systems are explainable and transparent, allowing users to understand how decisions are reached. Finally, the monitoring and maintenance processes, which are particularly critical in AI systems capable of learning from their use, must ensure adherence to established thresholds for acceptable performance, detecting and rectifying bias that may emerge as data distributions change over time, and protecting users against cyber threats and data breaches.

Global AI Governance must consider concentration of power.

Training large AI models, such as deep neural networks used in natural language processing or computer vision tasks, is a resource-intensive process in terms of hardware, software, human expertise, access to data, and time. It often requires a substantial investment in infrastructure and expertise to achieve state-of-the-art results. Currently, only a handful of very large technology companies, research institutions, and governmental organizations have the capacity to train AI models without relying on large cloud-based solution providers. This generates a dangerous concentration of power. Any global AI governance must address this concentration of power and prevent abusive and monopolistic use of the technology.

Global AI Governance should preserve the right to opt-out.

Article 15(b) of the International Covenant on Economic, Social and Cultural Rights provides the ‘right to benefit from scientific progress and its application’. All human beings, regardless of location, socio-economic status or vulnerability, have the right to opportunities to contribute to the development and benefit from the use of AI.

However, while we strongly support not leaving anyone behind in terms of opportunity and accessibility, ‘leaving no one behind’ should not become rationale for forcing universal adoption of AI systems, particularly regarding basic public services. The right does not require that individuals must use ‘scientific progress’, such as AI, even when the technology will arguably benefit the individual. In fact, as articulated by Committee on Economic, Social and Cultural Rights, “States parties must guarantee everyone the right to choose or refuse...with the full knowledge of the risks and benefits.”

Global AI Governance requires a new model.

Technology development evolves quickly and a global AI governance model must be agile enough to respond to the changing landscape. To help address this challenge, any governance body tasked with monitoring AI will need to include a robust forecasting unit, staffed with international experts capable of anticipating rapid algorithmic developments and new AI applications.

The international regime and organization for the prohibition of chemical weapons (OPCW) might be a helpful point of reference for the emerging global AI regulatory regime. Like AI, chemicals can be used not only as weapons, but for a broad variety of uses and they are ubiquitous across society in industry, agriculture, healthcare, manufacturing, and construction. For these reasons, the global regime for the prohibition of chemical weapons, which adopts a more holistic approach to the international regulation of chemicals, might provide an appropriate framework from which to draw inspiration.

The Center for the Study of Human Rights and Digital Technology¹ is a Paris- based research platform serving scholars from all disciplines who are working on questions of ethics, rights and technology. Founded in 2014 as a French association, the Center is a bilingual entity with the goal of facilitating conversations across European networks for scholars and practitioners with an interest in technology issues.

The Center’s research joins ethics, law and computer science in an attempt to understand more fully the dense, multidimensional nature of the digital revolution and how we are going to live with it.

Selected publications:

Roda C. (ed.) (2011) *Human Attention in Digital Environments*. Cambridge University Press; FP7 Project PRIPARE, Project Number: ICT-610613, \$1.1m, European Commission (2013-2015).

Roda, C., & Perry, S. (2014) “Mobile Phone Infrastructure Regulation in Europe: Scientific Challenges and Human Rights Protection.” *Environmental Science & Policy* 37: 204–14.

Perry, S. & Roda, C. (2014) Teaching Privacy by Design to Non-Technical Audiences. *Springer CCIS Series*, Vol. 470.

Perry, S.& Roda, C. (2014) Privacy-by-Design Curriculum. Selected by the European Union Agency for Network and Information Security (ENISA) for the *Roadmap for NIS Education Programmes in Europe*.

Perry, S. (2015) *Illusion Pixel*. Paris: Lemieux Editions.

Perry, S. & Roda, C. (2017) *Human Rights and Digital Technology: Digital Tightrope*. Palgrave Macmillan.

Doyle, W. & Roda, C. (eds.) (2019) *Communication in the Era of Attention Scarcity*. Palgrave Macmillan UK.

Roda, C. & Perry, S. (2021) “Learning in Lockdown: Teaching Human Rights Practice During the COVID-19 pandemic”. *Journal of Human Rights Practice*, vol. 13, Issue 3, 690–702.

Stahl, B.; Rodrigues, R.; Santiago, N.; Macnish, K. (2022) “A European Agency for Artificial Intelligence: Protecting fundamental rights and ethical values”, *Computer Law and Security Review*, 45.

Stahl, B.; Antoniou, J.; Bhalla, N.; Brooks, L.; Jansen, P.; Lindqvist, B.; Kirichenko, A; Marchal, S.; Rodrigues, R.; Santiago, N (2023) “A systematic review of artificial intelligence impact assessments”. *Artificial Intelligence Review*, 56.

Stahl, B.; Brooks, L.; Hatzakis, T.; Santiago, N.; Wright, D. (2023) “Exploring ethics and human rights in artificial intelligence – A Delphi study”, *Technological Forecasting and Social Change*, 191.

Roda, C.; Stone, R.; Kung, A. (forthcoming) “Standardization for high impact technologies: why citizens should care” in Ljubiša Bojić, Jörg Matthes, Damian Trilling, Simona Žikić (editors) *Emerging Technologies & Society: A Multidisciplinary Outlook*. Published by University of Belgrade, Institute for Philosophy and Social Theory and University of Vienna, Department of Communication.

¹ Centre d’études en Droits humains & Technologie numérique
Association régie par la loi du 1 juillet 1901 – SIRET 812 196 947 00012