Human Rights Law Centre

Inquiry on Adopting Artificial Intelligence

10 May 2024 / David Mejia-Canales

Human
Rights
Law
Centre

David Mejia-Canales
Senior Lawyer
Human Rights Law Centre
Level 17, 461 Bourke Street
Melbourne VIC 3000

T: + 61 3 8636 4450
F: + 61 3 8636 4455
E: david.mejia-canales@hrlc.org.au
W: www.hrlc.org.au

Via email: aicommittee.sen@aph.gov.au.

Human Rights Law Centre

The Human Rights Law Centre uses strategic legal action, policy solutions and advocacy to support people and communities to eliminate inequality and injustice and build a fairer, more compassionate Australia. We work in coalition with key partners, including community organisations, law firms and barristers, academics and experts, and international and domestic human rights organisations.

The Human Rights Law Centre acknowledges the people of the Kulin and Eora Nations, the traditional owners of the unceded land on which our offices sit, and the ongoing work of Aboriginal and Torres Strait Islander peoples, communities and organisations to unravel the injustices imposed on First Nations people since colonisation. We support the self-determination of Aboriginal and Torres Strait Islander peoples.

Follow us at http://twitter.com/humanrightsHRLC

Join us at www.facebook.com/HumanRightsLawCentreHRLC/

Contents

1. Introduction

In the rapidly evolving landscape of artificial intelligence (**AI**), the need for robust, yet flexible, regulation has become evident.

1.1 The regulation of AI must be underpinned by human rights principles

Governments worldwide are grappling with the complexities of integrating AI into our lives while balancing its benefits. In this swiftly evolving landscape, the need for robust yet adaptable regulation is clear.

AI technologies offer new economic prospects but also raise concerns about potential job displacements. They have the capacity to enhance environmental management and agricultural production, yet they also harbour the risk of propagating harmful misinformation campaigns that threaten social cohesion and democratic processes.

Furthermore, the impact of AI on public decision-making processes, which may be subject to judicial review, requires careful consideration to guarantee accountability and transparency. In discussions about the benefits of AI, it is crucial to highlight the essential role of human rights law and principles as the cornerstone of any regulatory framework.

Neglecting to address AI regulation through a human rights lens not only jeopardises individual liberties but also exacerbates societal disparities and undermines democratic governance.

Australia's absence of a comprehensive regulatory framework that is grounded in human rights law leaves a void that private entities will exploit. Allowing corporations to operate within this policy vacuum poses significant risks, particularly as profit-driven motives will likely prioritise lucrative AI applications over socially beneficial ones.

The opacity surrounding AI systems compounds these risks, hindering public understanding and oversight. Recent instances of malfunction or unfair decision-making highlight the urgent need for comprehensive regulatory mechanisms rooted in human rights law.

Government intervention is essential to ensure that AI advancements align with broader societal interests and ethical imperatives.

Without decisive action guided by human rights principles, there is a risk of deepening social divides and undermining the foundations of a just and inclusive society.

2. Human rights and AI regulation

Human rights law and principles are often overlooked in discussions about AI governance.

AI governance principles, crafted by companies, governments, civil society, and international organisations, frequently ignore human rights altogether.[1] Of those that do mention human rights, only a small fraction, approximately 15 percent, actually use them as a guiding framework.[2] Moreover, most national AI strategies do not deeply consider human rights at all.[3] This is unfortunate because human rights law provides established, clear and specific principles that hold international legitimacy.

Emerging from the aftermath of global conflict, the human rights movement was established with the explicit goal of preventing human misery and promoting peace. At the core of human rights law is the belief that its principles are not just crucial for averting catastrophe but also for ensuring that all people can live with dignity.[4]

Human rights law offers a robust framework that transcends borders, providing mechanisms for implementation, oversight, and accountability essential for effective governance in the AI era. Their importance lies in their capacity to guide ethical decision-making while fostering a society where everyone's rights are respected.

In discussions about AI regulation, it's crucial to understand that the focus should not be on whether human rights can or should apply to AI. Instead, the emphasis should be on recognising how the existing human rights framework is inherently relevant to AI governance.

2.1 Ethics and AI

In the discourse surrounding AI regulation, ethics often takes precedence over human rights principles.

Both ethics and human rights are integral to AI governance, each offering unique but complementary roles. While they might share an overarching aim of restraining state and corporate power and safeguarding individual interests, they employ different methodologies. It's imperative to recognise that neither can serve as a complete substitute for the other; rather, both must be seamlessly integrated into AI regulation endeavours.[5]

Ethics play a pivotal role in shaping and reinforcing regulation within the AI sphere. However, its philosophical underpinnings lead to a lack of standardised norms, resulting in a global absence of consensus on its precise applicability in AI governance. Relying solely on ethics could pose challenges for companies, public bodies, and individuals in understanding what their respective responsibilities are when it comes to AI development and deployment.[6]

Additionally, unlike human rights law, ethics lacks mechanisms to ensure accountability. Conversely, human rights offer a well-established legal framework that should be regarded as parallel to, rather than secondary to, ethics.

---

[1] Kate Jones, *AI Governance and human rights- resetting the relationship* (Chatham House International Law Programme Research Paper, January 2023) 10.
[2] Ibid.
[3] Ibid.
[4] Ibid.
[5] Kate Jones (n 1) 11-12.
[6] Ibid.

2.2 Human rights in the digital realm

Australia, as a party to the major international human rights treaties,[7] is required to guarantee and protect the rights enumerated in those treaties to all people in Australia's jurisdiction, including in the digital realm. The United Nations' Human Rights Council has issued resolutions requiring that 'the same rights that people have offline must also be protected online'.[8]

Businesses developing or deploying AI technologies do not have the same obligations to guarantee and protect human rights in the way that nation states do. However, international human rights law is not silent on the role of businesses in the protection and promotion of human rights.

The *United Nations' Guiding Principles on Business and Human Rights* provides a framework to ensure that the activities of all businesses, regardless of their size, corporate structure, ownership or location; are human rights compliant.[9]

The Principles assert that governments must take appropriate measures to ensure that companies and other non-state entities uphold human rights. Moreover, these principles emphasise that companies, including AI developers, have a duty to respect human rights in all their global operations, including through the implementation of due diligence and impact assessments.[10]

Companies' understanding of their human rights responsibilities varies, but there is a growing recognition that they must address human rights issues. This shift is driven by several factors, including evolving AI regulation, pressure from investors, community groups, and the public.[11]

---

[7] These are: *The International Covenant on Civil and Political Rights*; *The International Covenant on Economic, Social and Cultural Rights*; *The International Convention on the Elimination of All Forms of Racial Discrimination*; *The Convention on the Elimination of Discrimination against Women*; *The Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment*; *The Convention on the Rights of the Child*, and *The Convention on the Rights of Persons with Disabilities*.
[8] UN Human Rights Council Resolutions (2012-2018), *The promotion, protection and enjoyment of human rights on the Internet*, UN Doc A/HRC/RES/38/7 (5 July 2018), A/HRC/RES/32/13 (1 July 2016), A/HRC/RES/26/13 (26 June 2014), A/HRC/RES/20/8 (5 July 2012) as referenced in Kate Jones, *Online Disinformation and Political Discourse- Applying a Human Rights Framework*, (Chatham House International Law Programme Research Paper, November 2019) <https://www.chathamhouse.org/sites/default/files/2019-11-05-OnlineDisinformation-Human-Rights.pdf> 30.
[9] United Nations Office of Human Rights, *Guiding Principles on Business and Human Rights*, HR/PUB/11/04, 2011, <https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf>
[10] Kate Jones (n 1) 13.
[11] Ibid.

3. Adopting Artificial Intelligence

3.1 International approaches to mitigate AI risks

3.1.1 The European Union

On April 8, 2019, the European Union's (**EU**) High-Level Expert Group on AI introduced the *Ethics Guidelines for Trustworthy Artificial Intelligence*, which outline the principles necessary for AI to be considered trustworthy.[12]

The guidelines emphasise that for AI to be considered trustworthy it must be:

1) **Lawful**: AI respects all applicable laws and regulations, including international, regional and local human rights laws.
2) **Ethical**: AI respects ethical principles and values.
3) **Robust**: AI is robust, resilient and secure in both its technical dimensions while also accounting for the social environment it operates in.

The Guidelines outline the key requirements that AI systems should meet to ensure trustworthiness. These requirements encompass various aspects, including human agency and oversight, technical robustness, safety, privacy, data governance, transparency, non-discrimination, fairness, societal and environmental well-being, and accountability.[13]

Each requirement is accompanied by specific assessment criteria aimed at evaluating its implementation. These criteria include empowering human decision-making, fostering diversity and inclusivity, promoting environmental sustainability and ensuring accountability and redress mechanisms.[14]

In April 2021, the European Commission proposed the EU's first regulatory framework for AI, known as the *AI Act (***the Act**). The Act categorises AI systems based on their risk levels, with more regulation applied to higher-risk systems.[15] The Act aims to promote trustworthy AI by ensuring respect for fundamental rights, safety, and ethical principles while addressing risks posed by powerful AI models.

The regulatory framework defines four risk levels for AI systems:[16]

Unacceptable Risk: AI systems posing clear threats to safety, livelihoods, and rights will be banned, these include social scoring systems or even toys encouraging dangerous behaviour.

High Risk: This category includes AI used in critical infrastructure, education, the safety components of products, employment, public services, law enforcement, migration management, and democratic processes. High-risk AI systems face strict obligations before market entry, including risk assessment, data quality assurance, transparency, documentation, human oversight, and security measures.

Limited Risk: Refers to the risks associated with a lack of transparency in AI use. The Act introduces transparency obligations over limited risk applications to ensure humans are informed when they are interacting with AI so they can take an informed decision to continue to engage or not. For example, requiring that chatbots and AI-generated content are labelled as such.

---

[12] European Commission, *AI Ethics Guidelines for trustworthy AI* (Report, 8 April 2019) <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.
[13] Ibid.
[14] Ibid.
[15] European Commission, *AI Act* (Website, 6 May 2024) <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>.
[16] Ibid.

Minimal or No Risk: Allows for the free use of low-risk AI applications, such as those in video games or email spam filters, with minimal regulation.

The Act takes a forward-looking approach, designed to adapt to the rapid evolution of AI technology. It emphasises the need for AI applications to remain trustworthy even after they enter the market, requiring ongoing quality and risk management by providers to ensure compliance with regulations and laws, including human rights law.

To enforce and implement the Act, the European AI Office was established within the European Commission in February 2024. This office is tasked with overseeing enforcement, in collaboration with EU member states, with a focus on ensuring that AI technologies respect human rights and foster trust. Additionally, it promotes collaboration, innovation, and research in AI among various stakeholders and engages in international dialogue to align global AI governance standards.[17]

Following an agreement reached in December 2023 between the European Parliament and the Council of the EU, the AI Act is now in the process of formal adoption and translation. Once published in the *Official Journal*, it will enter into force 20 days later, with full applicability after two years.

The Commission has also launched the *AI Pact*, a voluntary initiative inviting AI developers to comply with key obligations of the AI Act ahead of time, aiming to facilitate a smooth transition to the new regulatory framework.[18]

3.1.2 Brazil

Brazil is considering *Bill 2338/2023*, which mirrors the EU's *AI Act* in its risk-based approach, tailoring obligations to the level of risk posed by AI systems.[19]

Like the EU Act, *Bill 2338/2023* prohibits systems with excessive risk, including those employing subliminal techniques to induce harmful behaviour and those used for social scoring by public authorities.

Penalties for violating this proposed law range from fines up to R$50,000,000.00 (AUD$ 14,940,532.27) per infraction or up to 2% of a companies' annual revenue, along with other penalties like the suspension of AI system development or supply.

3.1.3 Canada

In June 2022, Canada introduced the *Artificial Intelligence and Data Act* (**AIDA**) as part of the Canadian *Digital Charter Implementation Act*.[20]

Similar to the EU's AI Act, the AIDA proposes a risk-based approach, requiring high-impact AI systems to meet safety and human rights standards.

The Minister of Innovation, Science, and Industry would oversee the AIDA's enforcement, supported by a new AI and Data Commissioner. AIDA also introduces criminal provisions to prevent reckless AI use and ensures accountability for AI systems in international and interprovincial trade. The AIDA is under committee review in the Canadian House of Commons.

---

[17] Ibid.
[18] Ibid.
[19] Chase Young, Gabriel Mallare, Cella Kamarga, Yixuan Wu, *Regulation Comes to AI* (Website, 19 April 2024) <https://business.cornell.edu/hub/2024/04/19/regulation-comes-to-ai/>.
[20] Government of Canada, *The Artificial Intelligence and Data Act (AIDA)* Companion Document (Website, 13March 2023) <https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document>.

3.1.4 Australia

The University of Technology Sydney's Human Technology Institute has developed an outline for a model law addressing facial recognition technology (**FRT**). This initiative could assist in developing regulation for AI more broadly.[21]

The model law aims to restrict or even prohibit the development and deployment of FRT that poses risks to human rights while facilitating the use of FRT that aligns with Australia's liberal democratic principles. The model law would primarily target FRT developers and deployers operating within Australia.

Key objectives of the model law include upholding human rights, adopting a risk-based approach to classify FRT applications according to their potential harm, supporting compliance among developers and deployers, ensuring transparency in FRT usage, establishing effective oversight mechanisms through a well-resourced regulator, and providing avenues for accountability and redress for people affected by FRT misuse.[22]

Additionally, the model law seeks jurisdictional compatibility to ensure consistent application across all levels of government and adherence to global standards for international interoperability. By addressing these goals, the model law aims to navigate the complexities of FRT regulation in a manner that safeguards individual rights and fosters responsible technological innovation in Australia.

3.2 Australia should take a risk-based approach to AI regulation

The EU's approach, exemplified by the *Ethics Guidelines for Trustworthy Artificial Intelligence* and the *AI Act*, provides a robust framework for regulating AI systems based on their risk levels.

Brazil's proposed legislation, *Bill 2338/2023*, mirrors the EU's approach, demonstrating the international recognition of the importance of risk-based regulation. Similarly, Canada's AIDA proposes a risk-based approach, emphasising safety and human rights standards for high-impact AI systems.

Australia can ensure the responsible development and adoption of AI technologies by embracing a risk-based regulatory framework, such as the one proposed by the Human Technology Institute's model law for FRT. This approach fosters trust among individuals and businesses, promotes innovation, and encourages collaboration on a global scale.

[21] Nicholas Davis, Lauren Berry, Edward Santow, *Facial recognition technology towards a model law* (Report, September 2022).
[22] Ibid 10.

4. Recommendation

As Australians increasingly embrace AI technologies across various sectors, there is a growing awareness of the potential threats they pose if they are left unregulated.

While AI holds promise for enhancing efficiency and decision-making, its deployment also raises concerns regarding transparency, accountability, and the erosion of democratic principles. In light of this, the Human Rights Law Centre recommends the following:

4.1 The Parliament should consider adopting a risk-based approach to AI regulation

Australia should carefully consider adopting a risk-based approach to AI regulation, grounded in international human rights law and principles.