



Australian
Human Rights
Commission

Centring Human Rights in the Governance of Artificial Intelligence

Australian Human Rights Commission

Submission to the United Nations Office of the Secretary-General's
Envoy on Technology

30 September 2023

ABN 47 996 232 602
Level 3, 175 Pitt Street, Sydney NSW 2000
GPO Box 5218, Sydney NSW 2001
General enquiries 1300 369 711
Complaints info line 1300 656 419
TTY 1800 620 241

Australian Human Rights Commission
www.humanrights.gov.au

Contents

1	Commission introduction	3
2	Human rights risks of AI	4
2.1	<i>Privacy</i>	4
2.2	<i>AI interoperability</i>	4
2.3	<i>Automation bias</i>	5
2.4	<i>Algorithmic bias</i>.....	6
3	Human rights-centred artificial intelligence	6
3.1	<i>Regulatory responses</i>	6
3.2	<i>Non-regulatory measures</i>	8
4	Recommendations	8

1 Commission introduction

1. The Australian Human Rights Commission (Commission) welcomes the opportunity to make this submission to the United Nations Office of the Secretary-General's Envoy on Technology in response to the [Call for Papers on Global AI Governance](#) (Call for Papers) to assist the High-Level Advisory Body on Artificial Intelligence (Advisory Body).
2. As Australia's National Human Rights Institution, the Commission works towards a world in which human rights are respected, protected, and promoted. While the Commission has expertise and knowledge in human rights generally, relevant to the Call for Papers, it also has specific experience in considering the human rights issues in respect of technology.
3. This can be seen in the Human Rights and Technology Project, which was a three-year national investigation, that culminated in the release of the [Human Rights and Technology Project Final Report in 2021](#) (Final Report).
4. More recently the Commission, in partnership with the Actuaries Institute, published guidance on [Artificial Intelligence \(AI\) and discrimination in insurance pricing and underwriting](#).
5. In 2023, the Commission has continued its work to highlight issues concerning human rights and technology, including:
 - [The Need for Human Rights-centred AI](#): submission to the Department of Industry, Science and Resources.
 - [Utilising ethical AI in the Australian Education System](#): submission to the Standing Committee on Employment, Education and Training.
 - [Human Rights in the Digital Age](#): Global Digital Compact submission to the United Nations' Office of the Secretary-General's Envoy on Technology.
 - [Tackling Technology-facilitated Slavery](#): submission to the United Nations' Special Rapporteur on Slavery.
 - [Safeguarding the Right to Privacy](#): submission to the Attorney-General's Department in response to the Privacy Act Review Report 2022.
 - [Foreign Interference through Social Media](#): submission to the Senate Select Committee on Foreign Interference through Social Media.
 - [Privacy Risks in the Metaverse](#): submission to the Australian Competition and Consumer Commission.

6. This submission provides high-level recommendations for consideration by the Advisory Body. The Commission welcomes further opportunities to assist the Advisory Body in its work.

2 Human rights risks of AI

7. AI has the potential to improve our lives in meaningful ways. However, if it is not developed and deployed safely, it can also threaten human rights.

2.1 Privacy

8. The right to privacy is a cornerstone human right which also underpins freedoms of association, thought and expression, as well as freedom from discrimination.¹ The United Nations Human Rights Council has indicated that privacy is of increasing importance to everyday people in an age where:

... digital tools can be turned against them, exposing them to new forms of monitoring, profiling, and control.²

9. The operation of AI may facilitate and deepen privacy intrusions in new and concerning ways.³ AI products are trained on large data sets, which often include personal information – this incentivises a broad approach to collecting, storing, and processing huge amounts of data.⁴ Companies already aim to optimise services by collecting as much personal data as possible.⁵ The collection of data to train AI products will only heighten existing issues surrounding data collection.⁶
10. Despite the importance of the right to privacy, many private enterprises that build and deploy AI products have been reluctant to reveal details about the data used for training, or the data's provenance – which may raise issues of purchasing data and data scraping.⁷ The importance of these issues was highlighted by the [joint statement](#) on data scraping released by the Office of the Australian Information Commissioner and 11 of its international counterparts.

2.2 AI interoperability

11. It is important not to consider AI global governance in a silo. AI is an interoperable technology with the ability to integrate itself into the systems of several other technologies.
12. For example, AI has been used in combination with neurotechnologies to translate brain activity into words.⁸ In one experiment, AI was capable of

translating private thoughts into readable language by analysing fMRI scans, which measure the flow of blood to different regions of the brain.⁹

13. There are severe [human rights risks associated with neurotechnology](#) which must be considered in any examination of the global governance of AI. But it goes further than this – in investigating global governance, the Advisory Body needs to consider other technologies and their interoperability with AI.

2.3 Automation bias

14. The overreliance on AI-informed outcomes is a phenomenon known as ‘automation bias’, which is the

tendency to use automated cues as a heuristic replacement for vigilant information seeking and processing.¹⁰

15. In Australia, it is likely that automation bias played a role in the harms caused by the ‘robodebt’ scheme, in which an automated debt recovery system used an algorithm to identify undeclared income, resulting in a debt notice being automatically generated.¹¹

16. A parliamentary inquiry revealed that this process resulted in various inaccuracies which disproportionately affected people with pre-existing socioeconomic disadvantage and vulnerability.¹² A subsequent Royal Commission heard of the suffering this caused. One example was Kathleen Madgwick’s evidence to the Royal Commission of her son, Jarrad Madgwick, who had taken his own life just hours after he learned of a supposed \$2,000 Centrelink Robodebt.¹³

17. Since the release of the Royal Commission’s [report](#), the Australian Prime Minister, the Hon Anthony Albanese MP, stated that

The Robodebt scheme was a gross betrayal and a human tragedy, ... It pursued debt recovery against Australians who in many cases had no debt to pay. ... It was wrong. It was illegal. It should never have happened and it should never happen again.¹⁴

18. Although AI can be used in decision making with a ‘human-in-the-loop’, which may improve accountability and fairness, this approach in isolation is insufficient.¹⁵ Individuals who have oversight of decisions or processes informed by AI need greater training on the flaws of AI tools and must be encouraged to scrutinise AI-outcomes, especially where they can result in a significant consequence for an individual. It is imperative that the Advisory Body is not over-reliant on ‘human-in-the-loop’ recommendations without ensuring the training is in place to support those humans in decision-making.

2.4 Algorithmic bias

19. Algorithmic bias arises where an AI tool produces outputs that result in unfairness.¹⁶ Algorithmic bias can entrench unfairness and even result in unlawful discrimination.¹⁷
20. For instance, a 2019 study uncovered that a clinical algorithm used by many hospitals in the United States to determine which patients required extra medical care produced racial bias.¹⁸ The algorithm was trained on past data on healthcare spending, which reflects a trend whereby black patients have less income to spend on their healthcare as compared with white patients – a result of systemic wealth and income disparities.¹⁹ The algorithm’s outputs therefore reflected a discriminatory result of white patients requiring more medical care than black patients.²⁰
21. The Commission’s 2020 technical paper, [‘Using artificial intelligence to make decisions: Addressing the problem of algorithmic bias’](#), considers algorithmic bias in greater detail – highlighting technical considerations to avoid bias and discrimination.

3 Human rights-centred AI

22. It is of utmost importance that the Advisory Body have regard for human rights when considering the global governance of AI.
23. One example of how this is possible was outlined in the Commission’s [Final Report, which put forward](#) a model for AI regulation in Australia that had its basis in human rights protections.

Recommendation 1: The Advisory Body anchor its consideration of global governance of artificial intelligence in human rights principles.

3.1 Regulatory responses

24. To best protect human rights, AI governance requires binding regulation to create and apply safeguards to the development, training and deployment of AI products. The Commission encourages the Advisory Body to emphasise national regulation rather than self-regulatory models.

Recommendation 2: The Advisory Body should emphasise the regulation of artificial intelligence at the national level, rather than models based on self-regulation.

25. An appropriate regulatory framework may stimulate innovation and promote, rather than restrict, new activity by promoting stability and certainty in the regulatory environment.²¹
26. By creating a stable regulatory landscape for business to operate, AI innovation is better fostered by sensible regulation. Clear regulation and guidance can support businesses to ensure their products and services are free of unnecessary legal risk and can operate with certainty – to deliver profit without unduly exposing the company to unnecessary liability. The ambiguity presented by a lack of regulation can stifle innovation as businesses struggle to assess the legal and reputational risk that attaches to their products and services.
27. However, this does not mean that countries should enact AI-specific legislation without careful consideration. For example, Australia is actively working towards developing its approach to [AI regulation](#). Although this process is still underway, the Commission would highlight its [recommendations](#) here.
28. There are likely three pathways to regulating AI to mitigate notable human rights risks:
- create AI-specific legislation, analogous to other jurisdictions such as the EU;
 - reform and broaden existing regulation to ensure that it covers all applications of the technology; or
 - some combination of the two.
29. The Commission supports a combined approach to regulation. AI-specific legislation should be introduced in addition to reviewing and updating existing legal frameworks.
30. Any approach to global governance of AI should adopt a similar model – strengthening existing legislation and introducing AI-specific legislation to address identified shortcomings.

3.2 Non-regulatory measures

31. There are also non-regulatory responses to AI governance. The Commission has advocated for several non-legislative approaches, including the creation of a national AI Commissioner (an independent statutory body tasked with assisting the regulation and development of AI) and the need for human rights impact assessments to be completed on AI services and products.
32. The Commission is also highly supportive of the development of AI guidance for business. The Commission's own guidance on [artificial intelligence \(AI\) and discrimination in insurance pricing and underwriting](#) has been successful in assisting businesses to use AI.

Recommendation 3: The Advisory Body produces guidance and non-regulatory support for the governance of artificial intelligence – to complement binding forms of artificial intelligence governance.

4 Recommendations

33. The Commission makes the following recommendations.

Recommendation 1

The Advisory Body anchor its consideration of global governance of artificial intelligence in human rights principles.

Recommendation 2

The Advisory Body should emphasise the regulation of artificial intelligence at the national level, rather than models based on self-regulation.

Recommendation 3

The Advisory Body produces guidance and non-regulatory support for the governance of artificial intelligence – to complement binding forms of artificial intelligence governance.

Endnotes

¹ Office of the Australian Information Commissioner (OAIC), 'What is Privacy?' (Website) <<https://www.oaic.gov.au/privacy/your-privacy-rights/what-is-privacy>>.

- ² Office of the United Nations High Commissioner for Human Rights, *'The Right to Privacy in the Digital Age'* (Report, A/HRC/51/17, 04 August 2022) 2.
- ³ Office of the United Nations High Commissioner for Human Rights, *'The Right to Privacy In the Digital Age'* (Report, A/HRC/51/17, 04 August 2022) 4.
- ⁴ Office of the United Nations High Commissioner for Human Rights, *'The Right to Privacy in the Digital Age'* (Report, A/HRC/51/17, 04 August 2022) 4.
- ⁵ Wolfie Christl, *Corporate surveillance in everyday life* (Vienna, Cracked Lab – Institute for Critical Digital Culture, 2017).
- ⁶ Office of the United Nations High Commissioner for Human Rights, *'The Right to Privacy in the Digital Age'* (Report, A/HRC/51/17, 04 August 2022) 4.
- ⁷ G. Bell et al., *'Rapid Response Information Report: Generative AI - language models (LLMs) and multimodal foundation models (MFMs)'* (Australian Council of Learned Academies, Report, 24 March 2023) 12.
- ⁸ Oliver Whang, 'A.I. Is Getting Better at Mind-Reading', *New York Times* (01 May 2023) <<https://www.nytimes.com/2023/05/01/science/ai-speech-language.html>>.
- ⁹ See generally Jerry Tang et al., 'Semantic Reconstruction of Continuous Language from Non-invasive Brain Recordings' (2023) 26 *Nature Neuroscience*.
- ¹⁰ Max Schemmer, et al., *'On the Influence of Explainable AI on Automation Bias'* (Working Paper, 2022) 1 quoting Kathleen Mosier & Linda Skitka, 'Automation Use and Automation Bias' in Proceedings of the Human Factors and Ergonomics Society Annual Meeting (1999) 43(3) *SAGE Publications* 344–348.
- ¹¹ Commonwealth Ombudsman, *Centrelink's Automated Debt Raising and Recovery System: A report About the Department of Human Services' Online Compliance Intervention System for Debt Raising and Recovery* (April 2017).
- ¹² AHRC, Final Report (Final Report, 2021) 42.
- ¹³ Alexandria Utting, 'Kathleen Madgwick tells Robodebt royal commission about her son Jarrad and the damage the scheme caused' *ABC News* (Article, 10 March 2023) <<https://www.abc.net.au/news/2023-03-10/qld-robodebt-scheme-government-royal-commission-fraud/102027838>>.
- ¹⁴ Brett Worthington, 'What did the Robodebt royal commission find about the people who oversaw the scheme?' *ABC News* (article, 07 July 2023) <<https://www.abc.net.au/news/2023-07-07/political-reaction-to-robodebt-morrison-tudge-porter-robert/102575414>>.
- ¹⁵ G. Bell et al., *'Rapid Response Information Report: Generative AI - language models (LLMs) and multimodal foundation models (MFMs)'* (Australian Council of Learned Academies, Report, 24 March 2023) 12.
- ¹⁶ AHRC, Final Report (Final Report, 2021) 13.
- ¹⁷ AHRC, Final Report (Final Report, 2021) 13.
- ¹⁸ Crystal Grant, 'Algorithms are Making Decisions About Health Care, Which May Only Worsen Medical Racism' *American Civil Liberties Union* (Article, 3 October 2022) <<https://www.aclu.org/news/privacy-technology/algorithms-in-health-care-may-worsen-medical-racism>>.
- ¹⁹ Crystal Grant, 'Algorithms are Making Decisions About Health Care, Which May Only Worsen Medical Racism' *American Civil Liberties Union* (Article, 3 October 2022) <<https://www.aclu.org/news/privacy-technology/algorithms-in-health-care-may-worsen-medical-racism>>.

- ²⁰ Crystal Grant, 'Algorithms are Making Decisions About Health Care, Which May Only Worsen Medical Racism' *American Civil Liberties Union* (Article, 3 October 2022) <<https://www.aclu.org/news/privacy-technology/algorithms-in-health-care-may-worsen-medical-racism>>.
- ²¹ Michael Guihot & Lyria Bennett Moses, 'Artificial Intelligence, Robots and the Law' 2020 *LexisNexis Australia* 322.