

I, Robodebt

Submission to the Australian Human Rights Commission discussion paper on human rights and technology

Discussion paper



Director, Centre for Responsible Technology

March 2020

About The Australia Institute

The Australia Institute is an independent public policy think tank based in Canberra. It is funded by donations from philanthropic trusts and individuals and commissioned research. We barrack for ideas, not political parties or candidates. Since its launch in 1994, the Institute has carried out highly influential research on a broad range of economic, social and environmental issues.

About the Centre for Responsible Technology

The Australia Institute established the Centre for Responsible Technology to give people greater influence over the way technology is rapidly changing our world. The Centre will collaborate with academics, activists, civil society and business to shape policy and practice around network technology by raising public awareness about the broader impacts and implications of data-driven change and advocating policies that promote the common good.

Our philosophy

As we begin the 21st century, new dilemmas confront our society and our planet. Unprecedented levels of consumption co-exist with extreme poverty. Through new technology we are more connected than we have ever been, yet civic engagement is declining. Environmental neglect continues despite heightened ecological awareness. A better balance is urgently needed.

The Australia Institute's directors, staff and supporters represent a broad range of views and priorities. What unites us is a belief that through a combination of research and creativity we can promote new solutions and ways of thinking.

Our purpose - 'Research that matters'

The Institute publishes research that contributes to a more just, sustainable and peaceful society. Our goal is to gather, interpret and communicate evidence in order to both diagnose the problems we face and propose new solutions to tackle them.

The Institute is wholly independent and not affiliated with any other organisation. Donations to its Research Fund are tax deductible for the donor. Anyone wishing to donate can do so via the website at <https://www.tai.org.au> or by calling the Institute on 02 6130 0530. Our secure and user-friendly website allows donors to make either one-off or regular monthly donations and we encourage everyone who can to donate in this way as it assists our research in the most significant manner.

Level 1, Endeavour House, 1 Franklin St
Canberra, ACT 2601
Tel: (02) 61300530
Email: mail@tai.org.au
Website: www.tai.org.au
ISSN: 1836-9014

About the author

██████████ is the Director of the Centre for Responsible Technology. This report has been prepared with support from researchers from The Australia Institute, ██████████ and ██████████.

Summary

The Human Rights Commission's discussion paper on human rights and technology is an opportunity for Australians to rethink our relationship with technology and Artificial Intelligence (AI), and to address the "accountability gap" where technology escapes the regulation and scrutiny that exists in the "real" world.

The Commission has identified three goals for AI that could be expanded to serve as general principles for technology:

1. Technology should be used in ways that comply with human rights law
2. Technology should be used in ways that minimise harm
3. People should be accountable in how technology is used.

Existing human rights protections, like anti-discrimination laws, need to be enforced online and for "automated decisions" made by AI.

The use of AI and other technologies has so far been motivated by monetary concerns. Before a particular AI is adopted in Australia, its expected social impacts – on jobs, existing businesses and the vulnerable – should be assessed.

Automated decision-making should not be unaccountable decision-making. For each decision made by an AI, there should be a human identified as the person ultimately responsible for that decision.

Finally, the issues raised by AI and technology should provoke a more general discussion about whether government, the economy and society are driven by the right priorities. The size of the economy is a poor measure of human success and wellbeing. If human thriving was at the centre of decision-making, human rights concerns would be more readily addressed.

Introduction

The Australia Institute's Centre for Responsible Technology welcomes the opportunity to make a submission to the Australian Human Rights Commission's Human Rights and Technology discussion paper.

The discussion paper is an important contribution to thinking through profound change. By anchoring the government's response and responsibilities in the rule of law, it provides a compelling road map to manage change.

While regulation of Artificial Intelligence (AI) is a new topic in the Australian policy landscape, science fiction has been considering these issues for about 80 years. Isaac Asimov's Robot series centred around his Three Laws of Robotics:

First Law: A robot may not injure a human being or, through inaction, allow a human being to come to harm.

Second Law: A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.

Third Law: A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.¹

Interestingly, the Commission has proposed three key goals that in form and substance bear some resemblance to the sci-fi laws of robotics:

1. AI should be used in ways that comply with human rights law
2. AI should be used in ways that minimise harm
3. AI should be accountable in how it is used.

The Centre notes that AI is the aspect of technology that provoked the strongest community response.² However, the Centre believes that these goals are best stated as general goals for technology.

As with any technological advancement, it is *how* AI is used that determines whether human rights are threatened.

¹ See Wikipedia for extended discussion of Asimov's and others' work on this topic, https://en.wikipedia.org/wiki/Three_Laws_of_Robotics. In any discussion of Asimov and human rights, Asimov's own reputation for sexual harassment should also be mentioned. See Nevala-Lee (2020) *Asimov's empire, Asimov's wall*, <https://www.publicbooks.org/asimovs-empire-asimovs-wall/>

² Australian Human Rights Commission (2019) *Human Rights and Technology Discussion Paper*, p. 7, <https://tech.humanrights.gov.au/consultation>

Australia should require the responsible use of all technologies so as to comply with human rights law, to minimise harm, and to ensure real human accountability: including provisions for audits and appeals of decisions.

An AI can't be held accountable; it is motivated by programming and the ethics of its creator or user. The public program failure that has been “Robodebt” (formally, the “Online Compliance Intervention”) was primarily a failure of policy, not technology, as the Government of the day shifted the onus of proof for its own decision making to vulnerable people that rarely had any reasonable means of recourse.

Robodebt was a conscious decision to test a predictive algorithm around over-payment of on a vulnerable group of Australians. The impact of errors has been profound. One does not need to speculate on why the program was designed to identify people who had been subject of under-payments, or had not received payments they were entitled to. Of course, the reason is that the program was designed to generate an income stream (in alleged over-payments) for the government.

There is also an implicit learned helplessness at the moment about technology that is allowing an accountability gap to emerge. People sometimes see technology as a force outside of their control – but it is never “the computer's fault”. People should always be accountable for the decisions and outcomes of technology.

Accordingly, the Centre for Responsible Technology's first recommendation is to expand the three goals identified by the Commission to:

- 1. Technology should be used in ways that comply with human rights law**
- 2. Technology should be used in ways that minimise harm**
- 3. People should be accountable in how technology is used.**

Tech should respect existing laws

AI is incredibly complex and has the potential to fundamentally change the way we live, how our society operates and how our government makes decisions. Getting the settings right is vital.

Because technology is moving fast, the government is playing catch-up – but this is an opportunity to catch up by asserting the sovereignty of existing laws.

The Prime Minister says that “the rules that exist in the real world need to exist in the digital world”.³

By this logic, if anti-discrimination laws apply in the real world, technology cannot be permitted to work around those laws.

These legal frameworks give governments the power – and the responsibility – to ensure that technology that it uses – and that businesses operating in Australia use – are lawful.

Disruption that breaches the law is not innovative – it is illegal.

Like any citizen, the onus in a civil society needs to be on the technology developer to show they are acting legally, not wait for someone to identify a breach.

³ Martin & Karp (2019) *Consumer watchdog given new powers to crack down on Google and Facebook*, <https://www.theguardian.com/australia-news/2019/dec/12/consumer-watchdog-given-new-powers-to-crack-down-on-google-and-facebook>

Tech should be used in ways to minimise harm

Most AI projects being pursued by government and corporations are focussed on making or saving money.

This creates its own logic that determines what information that is collected and prioritised within any particular project.

With this the defining objective, rather than a broader public interest, the imperative to protect the public interest through a regulatory framework of checks and balances is essential.

When the money to be made in automating human functionality is so high, self-imposed codes of ethics are important, but not enough.

Before AIs are implemented in Australia, they should be reviewed – including requiring an AI Impact Statement that would require developers to articulate the impact of their product on jobs, existing businesses and on vulnerable sections of society

People should be accountable in how technology is used

The Centre for Responsible Technology submits that this human accountability needs to be both

1. abstract, that is explainable should someone appeal a decision, but also
2. concrete, that AI's should be authorised by an individual, much like political advertisements are.

Naming a responsible person connected to every AI will immediately generate a level of accountability that is currently missing.

This point is fundamental to the sort of society we are creating with Artificial Intelligence. Only with personal accountability will the principles in the Commission's report be fully realised.

Key points from the Commission Report

The Centre for Responsible Technology has identified a number of key points from the Commission that we strongly endorse.

Laws are needed to govern AI development, not simply ‘ethics’. Industry codes of ethics, standards and guidelines may have a place, but they are unenforceable and inevitably take second place to profit considerations. Existing legal frameworks must be enforced, and new frameworks introduced where needed.

Requiring humans to be involved in government decision-making would not just protect citizens’ interests, it would create a need for a new and skilled workforce to oversee automated decision-making in the public interest.

There should be a moratorium on the use of Facial Recognition Technology in significant decision-making until a legal framework is established (Proposal 11). This has two benefits: it protects citizens from ‘bad’ technology and it provides an incentive for the industry to support the development of the legal framework.

The hack of Australian start-up Clearview AI, which stores billions of photographs on its servers, in January 2020 reinforces the need for a moratorium on facial recognition technology until the multiple dangers it poses to human rights are addressed.⁴

Accountability. The report calls for anyone who deploys an AI-informed decision-making to be legally liable for the use of that system. Critically, this would mean that if an AI acted illegally or in a damaging way, there would be human accountability under both criminal and common law (Proposals 8–10).

Clear Processes. By creating its own transparent process to determine the benefits and costs of AI, government can set the ground rules for the commercial sector as well. The Centre supports a regulatory sandbox that could test AI-informed decision-making systems (Proposal 15).

Fairness-by-design. The Centre recommends that the government create a tool that would allow developers to undertake Human Rights Assessments of new products. This would set Australian best-practice and allow for the country to create a distinctly “Australian AI”

⁴ Centre for Responsible Technology (2020) *Massive Facial Recognition Hack Highlights Need to Call Pause*, <https://www.tai.org.au/content/massive-facial-recognition-hack-highlights-need-call-pause>

approach that could be differentiated from a surveillance AI approach used in other countries.

Review of existing AI. The public has a right to know what automated decision-making the government is already deploying or planning to deploy, and whether it complies with human rights and community expectations. A transparent review is urgently needed (Proposal 17).

Government influence. The government requiring AI best practice in its procurement policy would drive best practice across the industry (Proposal 18).

Oversight. We recommend an AI Safety Commissioner with standing powers to enforce laws, the “first model” envisaged by the Commission (Proposal 19).

Government should be driven by human success measures

The Centre for Responsible Technology's final recommendation is that this report should be implemented as part of a broader re-evaluation of how new technology should be embraced. The era of 'disruption', where products are trialled without consequence and 'breaking things' is celebrated, needs to end.

It is true that "you get what you measure", so why not systemically measure human outcomes? Not all human outcomes are financially measurable.

Alternative measures of human success include the concept of Gross National Happiness, a term coined by Bhutan's then king in the 1970s and influential on academic and popular thinking since then. Since Bhutan became a modern liberal democracy in 2008, its Constitution has required the state to promote the conditions that will enable the pursuit of Gross National Happiness.⁵ By measuring happiness, Bhutan places more value on the public programs, investment and behaviours that could promulgate it.

Similar measures include the Genuine Progress Indicator (GPI), which considers environmental, social and capital factors along with the narrow consumption or income factors measured by Gross Domestic Product. In 1997, Clive Hamilton and Hugh Saddler at The Australia Institute developed the Australian index of the Genuine Progress Indicator (GPI).⁶

Many countries have introduced human measures of success, profoundly changing the motivations and expectations of people, programs and organisations. Economic motivations subtly (or overtly) prioritise money over humans, which can lead to pursuit of short-term financial outcomes that often come at the expense of longer-term human outcomes.

New Zealand's Wellbeing Budget last year introduced "intergenerational wellbeing" measures of human, social, environmental and economic capitals. While in its early days, the Wellbeing Budget has begun to change behaviours and assumptions – including

⁵ For discussion of the history of the concept of Gross National Happiness, and the distinction between how it was applied in pre-democratic vs democratic Bhutan, see Correa (2017) *The History of Gross National Happiness*,

https://www.researchgate.net/publication/317151566_The_History_of_Gross_National_Happiness

⁶ Hamilton & Denniss (2000) *Tracking Well-being in Australia: The Genuine Progress Indicator 2000*, <https://www.tai.org.au/node/898>; Hamilton & Saddler (1997) *The Genuine Progress Indicator for Australia*, <https://www.tai.org.au/node/874>

encouraging ministers to collaborate and requiring them to consider wellbeing priorities in their funding proposals.⁷

The NSW Government's Human Services Outcomes Framework governs the design, delivery and evaluation of human services by government agencies and non-government organisations, using seven broad areas of quality of life measures: social and community, education and skills, empowerment, economic, safety, health, and home.⁸

This kind of framework, already developed and in place in the Australian context could be serve as a framework nationally for measuring public investment, programs, grants outcomes, taxation outcomes, social services, etc.

Without a national human outcomes measurement framework, there is no reference point for the most fundamental of questions: how is this technology contributing to the public interest?

Australians need to take time now to recalibrate how "success" is measured. Otherwise, more and more Australians will be left behind in what should be an era of surplus and human thriving.

⁷ Mintrom (2019) *New Zealand's wellbeing budget is a major policy innovation*, <https://www.themandarin.com.au/109476-new-zealands-wellbeing-budget-is-a-major-policy-innovation/>

⁸ NSW Government (2018) *The Human Services Outcomes Framework*, https://www.finance.nsw.gov.au/human_services

Conclusion

To guarantee the ethical use of technologies that maintains and protects human rights, human dignity and quality of life, humans must be both accountable and involved, and human measures of success are needed.

Isaac Asimov identified the Three Laws of Robotics in 1942, recognising almost 80 years ago that human wellbeing was at risk from increasing use of technology and needed to be central to its operating laws.

Getting these settings right is urgent. The broader “system” will continue to change and accelerate, so now is the time when measures of human success should be designed, implemented and valued. Measures of success, which inform business cases, grants, taxation and investment policy, should not be oblivious to the impact on humans.

Public Responses

Research conducted for the Australian Human Rights Commission by Essential Research shows there is an appetite for clear exposition of automated decisions.

- Just over half of Australians (56%) are aware that AI technology is being used by government agencies.
- Most Australians (57%) are somewhat or very uncomfortable with government agencies using AI technology to make automated decisions.
 - Discomfort increases with age, but even among those aged 18–34 more were uncomfortable (46%) than comfortable (41%).
- Seven in 10 Australians (71%) think it is very important that the government is able to provide an explanation when AI automated decisions are made. A further 17% think it is somewhat important.

METHODOLOGY

This report presents and analyses polling data from The Essential Report, a fortnightly online omnibus conducted by Qualtrics.

The omnibus usually delivers 1000+ interviews. In theory, with a sample of this size, there is 95 per cent certainty that the results are within 3 percentage points of what they would be if the entire population had been polled. However, this assumes random sampling, which, because of non-response and less than 100% population coverage cannot be achieved in practice. Furthermore, there are other possible sources of error in all polls including question wording and question order, interviewer bias (for telephone and face-to-face polls), response errors and weighting.

The online omnibus is live from the Wednesday night and closed on the following Sunday. Incentives are offered to participants. Essential Research uses the Statistical Package for the Social Sciences (SPSS) software to analyse the data. The data is weighted against Australian Bureau of Statistics (ABS) data. All Essential Research staff hold Australian Market and Social Research Society (AMRS) membership and are bound by professional codes of behaviour.

RESULTS

As far as you are aware, do you think that artificial intelligence technology is used in the following situations?

	Yes	No	Unsure
Facebook to determine the content displayed in each user's feed	61%	16%	24%
The Australian legal system to determine the appropriate length of sentences and/or fines for offences	28%	38%	34%
Government agencies such as Centrelink and the Australian Tax Office to analyse customer data and speed up processing time	56%	18%	26%
Hospitals to remotely monitor patients' health statistics (such as insulin levels)	46%	23%	31%

Some government agencies do use artificial intelligence technology to make decisions. When a government agency such as Centrelink or the Australian Tax Office makes a decision using artificial intelligence, rather than a human decision maker, this is called an automated decision.

How comfortable are you about government agencies using artificial intelligence technology to make automated decisions which can affect you?

	Total	Male	Female	18-34	35-54	55+
Very comfortable	7%	9%	5%	10%	7%	5%
Somewhat comfortable	25%	29%	21%	31%	25%	20%
Somewhat uncomfortable	33%	28%	36%	28%	35%	33%
Very uncomfortable	25%	23%	26%	17%	21%	35%
Unsure	10%	10%	11%	14%	11%	7%

If a Government agency uses AI to make an automated decision that affects you, how important is it that they are able to explain how the decision was reached?

	Total	Male	Female	18-34	35-54	55+
Very important	71%	68%	74%	54%	70%	87%
Somewhat important	17%	20%	14%	25%	19%	8%
Not very important	4%	4%	3%	7%	3%	1%
Not at all important	1%	2%	1%	2%	1%	0%
Unsure	7%	6%	8%	12%	7%	4%

All of the following are examples of how artificial intelligence technology is currently used by businesses and Government agencies. How appropriate or inappropriate do you find the use of artificial intelligence technology in the following situations?

Centrelink identifying benefit recipients that they believe have been overpaid, and issuing a repayment demand

	Total	Male	Female	18-34	35-54	55+
Very appropriate	20%	20%	20%	17%	19%	23%
Somewhat appropriate	24%	24%	24%	24%	23%	25%
Neither appropriate, nor inappropriate	16%	16%	16%	20%	15%	13%
Somewhat inappropriate	15%	13%	17%	17%	14%	13%
Very inappropriate	17%	19%	16%	11%	20%	19%
Unsure	8%	8%	9%	11%	8%	6%

Security companies carrying out facial recognition on security camera footage to find suspects in a large crowd

	Total	Male	Female	18-34	35-54	55+
Very appropriate	30%	25%	33%	30%	27%	32%
Somewhat appropriate	30%	29%	31%	29%	28%	33%
Neither appropriate, nor inappropriate	12%	13%	11%	12%	16%	9%
Somewhat inappropriate	9%	11%	7%	10%	8%	9%
Very inappropriate	11%	13%	9%	8%	14%	10%
Unsure	8%	8%	8%	11%	7%	6%

The Australian Tax Office using voice recognition to identify and verify people calling on the telephone

	Total	Male	Female	18-34	35-54	55+
Very appropriate	18%	18%	18%	19%	18%	17%
Somewhat appropriate	25%	25%	25%	25%	24%	27%
Neither appropriate, nor inappropriate	18%	20%	17%	23%	19%	14%
Somewhat inappropriate	14%	13%	16%	11%	14%	17%
Very inappropriate	15%	17%	13%	12%	16%	17%
Unsure	9%	7%	11%	11%	9%	7%

Recruitment companies automatically scanning CVs to only include some applications before a human employee reviews them

	<i>Total</i>	<i>Male</i>	<i>Female</i>	<i>18-34</i>	<i>35-54</i>	<i>55+</i>
Very appropriate	9%	10%	8%	12%	8%	7%
Somewhat appropriate	21%	26%	17%	24%	21%	20%
Neither appropriate, nor inappropriate	18%	17%	19%	23%	17%	16%
Somewhat inappropriate	21%	19%	24%	14%	23%	26%
Very inappropriate	21%	21%	21%	15%	23%	23%
Unsure	9%	8%	10%	12%	7%	8%

The Department of Health and Human Services collecting medical information about an individual from a range of private data sources

	<i>Total</i>	<i>Male</i>	<i>Female</i>	<i>18-34</i>	<i>35-54</i>	<i>55+</i>
Very appropriate	11%	13%	9%	13%	10%	10%
Somewhat appropriate	25%	26%	24%	24%	24%	26%
Neither appropriate, nor inappropriate	18%	16%	19%	18%	21%	14%
Somewhat inappropriate	18%	19%	18%	20%	18%	17%
Very inappropriate	19%	19%	19%	14%	18%	24%
Unsure	9%	8%	11%	12%	8%	8%