The Australian Human Rights Commission
The Human Rights and Technology Project Team
By email: tech@humanrights.gov.au

17 April 2020

Digital Rights Watch submission to the Australian Human Rights Commission

About DRW: Digital Rights Watch (DRW) is a nonprofit charity that works to ensure that Australians are equipped, empowered and enabled to uphold their digital rights: https://digitalrightswatch.org.au. For more information about this submission, contact:

███████████████████████████

Thematic responses to Proposals and Discussion Questions

Strategy and governance of policy making in emerging technologies

(Proposals 1-3, 19)

DRW supports these proposals. We think this has been a space that has lacked focus for lawmakers, and at times, been overly-politicised. Experts in this field have too often been ignored, and coordinated efforts at holding decision makers accountable have struggled to gain traction. Processes for developing best practice in respect of policies and regulation ought to involve users, especially vulnerable ones, as well as designers and civil society organisations.

DRW supports an ALRC inquiry, though we do think that steps can and should be taken now to protect people's rights. Any inquiry by the ALRC must be drafted with urgency, and ideally, have bi-partisan support to follow through on recommendations before commencement.

The terms of reference ought to include a consideration of the apportionment of liability for AI products, the rights of users in situations where the product does not function as expected, the rights and obligations associated with collecting and using data sets that contain personal information, and accountability for decisions.

DRW also strongly supports the creation of an AI Safety Commissioner. In addition to the suggested fields of work, DRW believes the office of the Commissioner should be able to conduct investigations, in response to complaints or on its own motion, with associated powers and issue fines, similar to the OAIC.

Cause of action: serious invasion of privacy

(Proposal 4)

A statutory cause of action for serious invasions of privacy is well overdue. The proposal has been subject to numerous inquiries and thoroughly considered by a significant number of stakeholders (indeed, the ongoing failure to act upon these recommendations highlights a discouraging attitude among policy makers towards the importance of public participation). Any cause of action ought to be broad enough to cover private actors as well as those in government. It is also important that the cause of action made adequate provision for deterrence, rather than just being compensatory in nature. Like wage theft, there are instances where the harm associated with a data breach can be relatively small for an individual, but very widespread for a collective, such that the law needs to demonstrate a strong deterrence capability to avoid non-compliance. Existing mechanisms for group and representative actions

(in the *Privacy Act* and generally in various court legislations) should be available for those affected.

Legal rights and remedies: AI-informed decision making

(Proposals 5-10)

DRW supports the requirement to inform individuals when AI is materially used in a decision about them that affect their rights. It is important to avoid such a requirement becoming a standard form disclosure that is rendered meaningless (like privacy policies). As such, it ought to be in plain language, and published prominently. Ideally, the drafting of such a statement would be guided by a non-exhaustive statutory set of criteria for considerations that ought to be taken into account. It ought to also indicate how such a decision could be challenged. This requirement ought to be imposed on the private sector, not just the government sector.

DRW generally supports a requirement for the government to conduct a cost/benefit analysis before deploying these systems. However, we think this has the potential to be insufficient, as there are powerful cost incentives to make greater use of AI-informed decision making, even if they come at the expense of rights. Any cost benefit analysis stage should include a human rights impact analysis and funded consultation, rather than relying on underfunded civil society organisations to fill the gaps. DRW also supports a prioritisation of open source software and accountability in respect of data sources on the part of government. The government should not be using AI systems that have been trained on data collected in invasive ways, and that are tainted by biases and gaps that have not been properly identified and addressed in the design stage.

DRW generally supports the introduction of legislation regarding the explainability of AI-informed decision making. However, we think there are some critical aspects that ought to be included in this regime.

- Any explanation of such a decision should indicate the data sources relied upon to train any automated process, how it was collected and by who or what;
- Any explanation of such a decision should also provide an indication of how the AI has been tested and designed to limit biases;
- Any decision ought to be subject to accessible appeal rights.

DRW also supports a rebuttable presumption that the legal person who deploys an AI-informed decision-making system is liable for the use of the system.

It is important to find ways to ensure that upon any appeal, the individual's rights are respected until the dispute is resolved. Unless this is the case, there is a real risk of creating incentives among government decision makers to use poorly designed AI to assist decision making because it is cheap, and the costs are defrayed onto individuals who bear the cost of overturning them.

To avoid this, policy makers should consider including a deeming provision such that, upon appeal, the decision is deemed to have been made in the way most favourable to the individual concerned until the appeal is resolved. This ought to be accompanied by a prohibition on requiring the individual to return any benefit gained as a result of this deeming provision, even if the appeal is unsuccessful. In the case of robodebt, by way of example, it was many years before the faulty decision making was reversed by the court, causing untold human suffering and allowing a significant cost saving to the government. Reversing this incentive requires strong presumptions in favour of the individual at all stages.

It is important that policy makers consider ways to avoid lengthy, drawn out appeal processes. This could involve a requirement that, upon appeal, a human official, with authority to make (or remake) the decision, be required to do so, with reasons, within a short period of time (say 7 days). It is critical that government departments seeking to rely on AI for the purposes of making decisions are appropriately resourced to manage a workload that may arise from this kind of technology. This kind of resourcing should be factored into any cost benefit analysis (whether formal or informal).

DRW supports the adoption of the precautionary principle in relation to AI-informed decision-making systems, and agrees they should not be deployed where they could infringe on human rights.

Further research into this field is desperately needed. Academic research will be vitally important to creating rights-respecting policy, and DRW supports institutions like the Australian Research Council Centre of Excellence for Automated Decision-Making and Society conducting this research. Civil society organisations should also be involved in the process of engaging communities and individuals, and centres of expertise should use civil society as a conduit to the broader public where possible.

Frameworks and tools for government: AI-informed decision making

(Proposals 12-18)

DRW supports the creation of standards relating to AI-informed decision making that incorporate guidance on human rights compliance. DRW also agrees that to ensure such standards are fit for purpose and effective, Australia needs a broad task force to assist in mapping out the ways in which human rights can be embedded into all stages of the design process. This ought to include input from vulnerable users, independent experts and a diverse range of civil society organisations. It will also be important to consult with designers and engineers, so that any concept recognises the pressures and potential of the design process.

DRW agrees that Australia needs a human rights impact assessment tool for AI-informed decision making. This is to ensure that operational impacts of AI-informed decision making can

be properly scrutinised and mistakes can be identified. Again, to create an effective tool will require diversity of input from experts, users, designers and civil society.

DRW agrees that a regulatory sandbox would be useful, for educational purposes, but such an approach must avoid creating a situation where vulnerable people are experimented on by policy makers.

Facial recognition

(Proposal 11)

DRW supports a legal moratorium on the use of facial recognition technology in decision making.

However, DRW would recommend specific cases of facial recognition technology be classed separately. For example, the use of facial recognition technologies to co-validate identity based on facial photographs that have already been provided by the subject should be treated differently to the use of facial recognition technologies by law enforcement that scan video streams to identify or short-list individuals that may be wanted by authorities. In this second scenario, DRW calls for a complete ban on using facial recognition technologies by law enforcement agencies or the use of facial recognition technologies to aid mass public surveillance.

The development of a regulatory regime ought to include extensive public consultation, especially with organisations that represent groups most at risk of poorly designed technology of this nature.

AI Safety Commissioner

(Proposal 19)

DRW supports this proposal.

'Human rights by design' and education and vocational training

(Proposals 26-27)

DRW supports greater emphasis on 'human rights by design' for educational disciplines associated with technological development, for vulnerable users as well as users with disabilities. We note the work for the Association of Computer Mechanics, as well as many other similar organisations, that are doing important work in this space.

It is critical that other disciplines are included in this initiative, especially if the field makes use of big data, and trains machines using that data. In particular, DRW believes such programs

should be part of the curriculum in schools of government, schools of business and schools of medicine. Cross disciplinary collaboration will be vital to the success of these programs.

Disability and technology

(Proposals 20-28)

DRW supports the proposals in relation to improving accessibility of technology for people with disabilities at all stages of design.