

17 March 2020

Australian Human Rights Commission
By webform tech.humanrights.gov.au/consultation

Dear Commission,

AHRC Discussion Paper on Human Rights and Technology

About us

The Allens Hub for Technology, Law and Innovation ('the Allens Hub') is an independent community of scholars based at UNSW Sydney. As a partnership between Allens and UNSW Law, the Allens Hub aims to add depth to research on the diverse interactions among technology, law, and society. The partnership enriches academic and policy debates and drives considered reform of law and practice through engagement with the legal profession, the judiciary, government, industry, civil society and the broader community. More information about the Allens Hub can be found at <http://www.allenshub.unsw.edu.au/>. For this submission, we have partnered with Andrew Ray and Bridie Adams, law students from the ANU College of Law and researchers at the National Judicial College of Australia. Our submissions reflect our views as researchers and are not an institutional position of UNSW, the ANU, the NJCA or Allens. We welcome the opportunity to submit to the review, focusing in particular on Proposals 1 and 11 and Questions A-D.

Recommendations

Our main recommendations are as follows:

1. consider the human rights task as part of a broader law reform project, drawing lessons from the past particularly as to methodology in the context of public engagement;
2. define scope by reference to clear terminology that captures areas of concern better than "artificial intelligence" – specific questions that arise, including the possibility of restrictions or regulations, should be analysed contextually;
3. encourage research into improving AI-informed decision-making through amending the priorities of existing Australian Research Council programs, as opposed to a rebuttable presumption that decisions made by AI-informed decision-making systems are unlawful if appropriate reasons are not provided;
4. a broad approach should be taken in assessing the legality of AI-informed decisions rather than focussing on technical information;
5. strengthen existing freedom of information laws to promote release of information concerning AI-informed government decision-making by relaxing exemptions available to government departments and agencies;
6. reconsideration of a decision should be made by humans in the event of a challenge to the original decision made through an AI-informed process;
7. human involvement in AI-informed decision-making should be scaled based on the context and potential impact of the decision; and

A joint initiative of

Allens > < Linklaters



8. the use of AI facial recognition technology by law enforcement agencies should be subject to a moratorium until larger issues can be resolved.

Proposal 1: Strategy on New and Emerging Technologies: Learning from History

We endorse the proposal in Chapter 3 for a national strategy for the protection of human rights in the development and use of new and emerging technologies; indeed we would go further. We would argue that human rights are *one* set of considerations that need to be contemplated in law reform and regulatory action in the context of technological change. In addition to the need to align new technologies with social values (which includes internationally recognised human rights, but may go further), legal and regulatory frameworks will need to be adjusted to resolve uncertainties in the application of existing rules, ensure that rules are not over- or under- inclusive with respect to activities facilitated by new technologies and make changes where existing laws or regulatory approaches are no longer justified, no longer cost-effective, or no longer applicable.¹ The expertise required for this broader task is similar to the expertise required for the narrower one – a combination of legal (including human rights) and technical (across rapidly evolving fields of technology). It would therefore make sense to combine these functions in a horizon scanning agency, rather than separate the human rights role from the broader law reform task.

It is worth noting that there much to learn from history here, in particular the role played historically by the National Enabling Technology Strategy – Science and Technology Engagement Pathways group. This group developed a strategy for community consultation to feed into co-ordination of federal, state and territory activities around the health, safety and environmental implications of emerging technologies. That focus was different to what we are proposing, which is broader, but the consultation strategy remains a thoughtful and considered approach to involving the broader public in policy development. In particular, principles such as commitment and integrity, clarity of scope and objectives, inclusiveness, good process, quality information and knowledge sharing, dialogue and open discussion and impact on decision-making are particularly crucial in the context of community engagement with the issues being considered by the Commission.

It should be noted that any such agency would not, as the Commission points out, be about “regulating an entire technology, as a technology”. Instead, it should focus on ensuring that legal and regulatory frameworks are well-adapted to an evolving socio-technical landscape, curating it in ways that ensure alignment with human rights and other values of importance to Australians.²

Question A: “AI-informed decision making” and the challenge of regulatory scope

We are concerned with the proposed scope of the recommendations in Chapter 5, in particular the focus on ‘AI-informed decision-making’. AI is defined on page 59 as “broadly speaking...the range of technologies exhibiting some characteristics of human intelligence’. We believe this is an unhelpful scope for a variety of reasons:

- It is both over- and under- inclusive. It includes calculators because, when I divide two large numbers, I use ‘intelligence’ to do so. It excludes Robodebt (an example repeatedly cited in the Commission’s report) because that system aligned more with human mismanagement and hubris than with ‘human intelligence’. It also excludes the use of data-driven inferencing where the methodology is humans doing a statistical regression (say) rather than a machine using a random forest algorithm – even if the goal of both is to identify correlations for discriminatory action. Depending on the particular system/smart contract used, this definition may be broad enough to capture automated processes deployed under smart contracts in the future. This is likely broader in scope than the AHRC has intended.
- The problem with the term ‘intelligence’ is that it is an inherently positive frame (so regulation of AI-informed decision making does not necessarily include all data-driven decision-making or all automated decision-making). It is not the simulation of a positive trait (intelligence) that calls for human rights protection, but the methods of decision-making used in practice in circumstances that have legal or similarly significant effect.

¹ Lyria Bennett Moses, ‘Recurring Dilemmas: The Law’s Race to Keep Up with Technological Change’ (2007) 7 *University of Illinois Journal of Law, Technology and Policy* 239.

² Lyria Bennett Moses, ‘Regulating in the Face of Sociotechnical Change’ in Roger Brownsword, Eloise Scotford and Karen Yeung (eds), *The Oxford Handbook of Law, Regulation and Technology* (Oxford University Press, 2017) 573.

- The term is anthropomorphic. Machines do not ‘think’ in the same way that humans do, they use different techniques to perform similar tasks. Using a human trait like ‘intelligence’ as a metric is unhelpful at classifying a set of technologies. What is important is what tool is used, how it works, and whether it is useful and appropriate, not whether it simulates a human trait.
- Rather than attempting to define a technology, we believe it is more useful to consider how changes in socio-technical practice align with human rights (and other) values. This changes the nature of the question being asked from “how should we ensure that AI-informed decision making aligns with human rights” to “how can we ensure that human rights are protected given changes in how decisions are made, including the increasing use of data-driven inferencing and increasing reliance on automated systems”. One can then point to AI as a field of research that has had an impact on these broader socio-technical changes, but the inquiry is not limited to better governance of a peculiarly defined set of techniques.
- The methods of decision-making will continue to evolve, and the techniques used might not always be classified by computer scientists and others as AI. The question of whether something is or is not AI should be beside the point – if it changes how decisions are made in ways that impact on human rights (or other values), it needs to be investigated by the apparatus proposed in the Commission’s Discussion Paper. While most computer scientists would not classify the Robodebt system as AI, given its widespread impact³ it is precisely the kind of thing that should be captured by the Commission’s proposals.
- Requirements such as ‘reasons for decisions’ exist in particular contexts, such as administrative decisions⁴ and not in others, such as who a person decides to marry. There is no reason why the use of technology (including AI) should change the category of decisions for which explanations are required. The question is not *whether* AI-informed decisions need to be explained but *how* an explanation can be generated where required for a particular purpose. This is a highly technical area that could be the subject of ISO/Australian standards.

Generally speaking, in any law reform or regulatory task, we believe that what is important is context. This includes both the technological context (for example, the use of machine learning or other data-driven inferencing techniques) and the broader context (such as whether a decision is made in the public or private sector). It is this context that should determine any legal or regulatory response. Some technologies should perhaps be banned outright (the argument has been made in relation to fully autonomous weapons systems), while others should not be used in particular contexts (such as the use of open-ended data-driven inferencing in sentencing). Facial recognition systems are discussed further below.

While we agree that there can be one entity assigned with managing law reform in response to a set of socio-technical changes, constraining this entity to an awkwardly defined technological context (such as AI-informed decision-making) is unhelpful. A broader mandate will allow for projects with narrower legal and technological foci such as human rights or discrimination law, on the legal side, or machine learning, on the technical side. Prohibition or regulation of particular technological practices, where appropriate, would be included in the broad mandate, as would a project such as reform of discrimination law to encompass data-driven inferencing (whether using machine learning or statistical methods).

Questions B and C: Statements of Reasons and Assessing the “lawfulness” of an AI-informed decision-making system

We are concerned with the focus on technical information in assessing the lawfulness of an AI-informed decision-making system contained within Chapter 6 of the report. We are also concerned about the potential impact of the creation of rebuttable presumption that a decision was not lawfully made where a lack of a reasonable explanation is provided.

- As outlined above, there are potential challenges in generating statements of reasons for decisions made by automated systems, especially where the systems involve deep learning techniques such as neural networks. We therefore welcome the AHRC’s recommendation concerning prioritising research on “how to design AI-informed decision-making systems [that can] provide a reasonable explanation to individuals”. It is however in our view more desirable to encourage this research through amending the priorities of existing Australian Research Council programs.

³ Commonwealth Ombudsman, ‘Centrelink’s automated debt raising and recovery system’ (Report No 2, Commonwealth Ombudsman, April 2017), see especially 8.

⁴ See, eg, *Administrative Decisions Judicial Review Act 1977* (Cth) s 13.

- We do not agree that there should be a rebuttable presumption that decisions made by AI-informed decision-making systems are unlawful if appropriate reasons are not provided. In particular, we believe that context is crucial, particularly as to whether a decision is made in the public or private sector. Further, it is not clear what the constitutional basis of a law rendering private decisions unlawful would be. We also note that this would be a significant change to the existing law so that the full merits and effects of the proposal will need to be considered.
- In line with our comments above about the importance of context, it is likely that given the state of current technology, some tools will be inappropriate for particular kinds of decision-making due to a requirement that reasons be given (either at the time of the decision or upon exercise of an appeal right). Courts will look unfavourably upon decisions made in some circumstances without clear and understandable statements of reasons. In the context of decisions where reasons are already required, it is unclear what the creation of the proposed presumption will add. There may however be scope for government procurement processes to require more open and transparent access to underlying software where government needs to be able to be transparent about the software/processes it has followed. We note that in the private sector, it is rare that reasons need to be given at all, and where reasons are not required it is unclear why AI-informed decisions ought to be treated differently.
- In the case of deep learning systems, algorithms identify rules and patterns from training data. Given our existing inability to understand how and why some systems make decisions based on these patterns, it is likely that experts will be unable to assess whether such systems are basing their decisions on the basis of legislatively approved rules and criteria. Where systems use other data inputs or do not follow legislative rules perfectly, they play a policy role that was not authorised in the legislation itself. alter legislatively approved rules or be relied on by government to the detriment of those affected by decisions.⁵ This requirement to publish policies that may detrimentally affect individuals may already require the publication of algorithms and underlying data to enable experts to assess whether they are altering legislative rules. Any law reform in this area should take account of the role played by existing rules in offering reasonable, technology neutral protections.
- Additionally, the creation of a presumption likely violates the principle of technological neutrality, in that it would create special rights for individuals seeking judicial or merits review where they were challenging decisions that fell under the regime. This could encourage individuals to challenge decisions under this scheme, even if the link to technology was tangential.
- It is unclear whether the Commission's proposal regarding explainability of decisions (proposal 7) applies only to government decision-making or also to non-government bodies. For the purposes of the following we have assumed that given the focus on both throughout the Chapter, both are intended to be captured.
- We are concerned about the Commission's focus on technical information in order to assess the legality of a decision for several reasons:
 - Technical information, especially how AI-systems generate decisions and the training data used by those systems, could be personal information (for the purposes of the *Privacy Act*) or may be kept confidential for commercial reasons. Release of personal information would generally be contrary to the Australian Privacy Principles.
 - This approach would also not accord with the general principle of technological neutrality and could preference individuals affected by decisions captured by this law due to the technology used to make the relevant decision. This could have unintended effects, including making Australian governments and private actors overly apprehensive about using AI-informed systems and unnecessarily encouraging individuals to argue that a particular decision should fall within the proposed statute. A potential alternative, at least in the case of government decision-making, would be to strengthen existing freedom of information laws by relaxing the exemptions available to government departments and agencies. These amendments could, for example, reduce the number of exemptions agencies can rely on to deny releasing information,⁶ further promote release of information as the primary objective of the act, or strengthen government powers to compel agencies to release information in senate

⁵ For discussion, see Andrew Ray, 'Implications of the future use of machine learning in complex government decision-making in Australia' (2020) *ANU Journal of Law and Technology* (forthcoming). See also *Freedom of Information Act 1982* (Cth) ss 8-10. See especially s 10(2) which requires that a person is not 'subjected to any prejudice because of the application ... of any rule, guideline or practice in unpublished information, if [they] ... could lawfully have avoided that prejudice had [they] ... been aware of the unpublished information'.

⁶ In particular, agencies have relied upon the exemption of information that is commercially valuable as a justification not to release the programs/code used for certain automated decision-making processes.

.....

or other government enquiries.

- It is unclear whether technical information concerning the algorithm will actually assist applicants in assessing the legality of the AI-informed decision, and may instead simply increase the costs involved in challenging a decision (as technical experts will be needed on both sides). For example, while data used in training an algorithm *may* be helpful in some cases, for others it would simply add to the complexity of the court challenge and may present further problems for privacy .
- Statements of reasons (as noted above) are already required in circumstances that individuals are seeking to challenge government decisions. While these statements do not necessarily require the release of technical information, the statements must be understandable by the person seeking it and give them a clear understanding about why the decision was made.⁷ It is unclear whether these requirements are currently insufficient in relation to government decision-making and at a minimum this matter should be referred to the ALRC for further consideration.

Transparency requirements should take account of context. Not all decisions, particularly those made in the private sector, need to be rendered transparent. In government, there are a variety of existing transparency requirements in freedom of information laws and administrative law that can be extended in technology-neutral ways to deal with new issues associated with AI-informed decision-making. Government procurement practices could also be modified to consider using open source software in contexts where transparency and accountability requirements run counter to commercial confidentiality clauses or freedom of information laws could narrow the circumstances in which commercial confidentiality prevents disclosure.

Proposal 11: The Australian government should introduce a legal moratorium on the use of facial recognition technology in decision-making that has a legal, or similarly significant, effect for individuals, until an appropriate legal framework has been put in place.

Proposal 11 has taken on added importance given the rumoured use of Clearview AI facial recognition system by Australian law enforcement bodies such as the AFP and state-based police.⁸ The use of technology for law enforcement and other purposes without a legislative framework has occurred in the past, for instance when law enforcement and other agencies accessed location data prior to 2007 despite there being no legislative regime for access, and, will likely occur in the future if no steps are taken to provide a sufficient legislative/regulatory/ethical framework to prevent such actions.⁹ However, given the wide-ranging ramifications for the use of facial recognition AI and the number of organisations worldwide who hold grave concerns over the use of this technology,¹⁰ we agree with a legal moratorium being implemented until issues of transparency and sufficient safeguards can be addressed. We note the Parliamentary Joint Committee on Intelligence and Security’s (‘PJCIS’) recommendations in rejecting the Identity Matching Services Bill 2019 (Cth) and the Australian Passport Amendment (Identity Matching Services) Bill 2019 (Cth). The PJCIS recommended the Bills be re-drafted based on principles of privacy, transparency and robust safeguards, the provision of Parliamentary oversight, reasonable, proportionate and transparent functionality, the provision of annual reporting and detailed obligations of participants to the scheme.

Comparing the principles outlined by the PJCIS and the possibility that Clearview AI is already being utilised by state and federal law enforcement agencies without a legislative framework heightens concerns. The PJCIS expressed its concerns on the use of Clearview at

⁷ *Re Palmer and Minister for the Capital Territory* (1978) 23 ALR 196, 206, 209.

⁸ ‘Leaked Document Shows Australian Police Use Creepy Clearview AI Facial Recognition Software’, *Gizmodo Australia* (29 February 2020) <<https://www.gizmodo.com.au/2020/02/australian-police-clearview-ai-facial-recognition/>>.

⁹ *Telecommunications (Interception and Access) Amendment Bill 2007*.

¹⁰ ‘Facial Recognition: It’s Time for Action’, *Microsoft on the Issues* (6 December 2018) <<https://blogs.microsoft.com/on-the-issues/2018/12/06/facial-recognition-its-time-for-action/>> (‘*Facial Recognition*’); Sigal Samuel, ‘The Growing Backlash against Facial Recognition Tech’, *Vox* (27 April 2019) <<https://www.vox.com/future-perfect/2019/4/27/18518598/ai-facial-recognition-ban-apple-amazon-microsoft>>.

its hearings on the metadata retention scheme, posing questions to the AFP on its use of the product. These questions were taken on notice.¹¹

Further compounding issues, Clearview's client database was recently hacked, and it holds 3 billion photographs scraped from social media platforms, some of whom have protested the use of images from their platforms for the development of facial recognition systems. Scraped photographs may leave Clearview exposed to copyright issues and breaches of contractual agreements with platforms such as Facebook, Google and LinkedIn. It is also not known how accurate the Clearview system is leaving open the possibility that decisions such as the arrest of an individual based on AI could be flawed.¹² The Data to Decisions Co-operative Research Centre wrote a series of reports on the use of "open source" data, exploring privacy, copyright, contract and other issues which we could make available to the Commission on request.

Question D: How should Australian law require or encourage the intervention by human decision-makers in the process of AI-informed decision making?

As indicated above, we believe context is crucial in determining the circumstances in which human decision-makers should be required to intervene. If a person wishes to use AI-informed decision making to decide whom they should marry, that might be ridiculous, but it should not be an area in which the law should intervene

In the case of AI-informed (or automated) government decision-making, where a decision is reviewed or successfully challenged and must be remade, we believe that reconsideration should be undertaken by a human decision-maker. This, alongside existing laws requiring the generation of statements of reasons, should generally encourage government to design its systems in such a way that they can be reviewed and held accountable by human decision-makers. Without the requirement for a decision to be remade by a human decision-maker, given the consistency of automated decision-making processes, we would expect the same decision to simply be remade.

If laws are introduced to require human involvement in some kinds of decisions, the level of involvement required should scale based on the potential impact of the decision in the relevant context. For example, in the case of AI-systems that are making targeting or firing decisions in combat,¹³ there is clearly the need for significant human control and oversight as those decisions are effectively non-revocable. Conversely, decisions with much lesser impact, such as the decision to issue a parking fine or speeding ticket, may require no human involvement in the initial stage – provided the review processes in place are stringent enough. A scaled oversight approach has the benefit of encouraging efficiency and acknowledging that if strong enough review rights are present (and the decisions are capable of being reviewed due to the generation of sufficient statements of reasons) then businesses and government should not unnecessarily be burdened with the requirement to involve human decision-makers in-the-loop.

Yours sincerely,

██████████, Director

██████████

██████████

██████████

¹¹ Canberra Commonwealth Parliament; Parliament House, 'Public Hearings'

<https://www.aph.gov.au/Parliamentary_Business/Committees/Joint/Intelligence_and_Security/Dataretentionregime/Public_Hearings>.

¹² Jake Goldenfein, 'Australian Police Are Using the Clearview AI Facial Recognition System with No Accountability', *The Conversation*

<<http://theconversation.com/australian-police-are-using-the-clearview-ai-facial-recognition-system-with-no-accountability-132667>>.

¹³ There is no evidence that systems are currently used with this level of technology, however some automated weapons systems have been deployed in limited circumstances.

.....