

18 March 2019

Mr Edward Santow  
Human Rights Commissioner  
Human Rights and Technology Project  
Australian Human Rights Commission  
GPO Box 5218  
SYDNEY NSW 2001



By email: [tech@humanrights.gov.au](mailto:tech@humanrights.gov.au)

Dear Commissioner

### **Artificial Intelligence: Governance and Leadership**

The Law Council welcomes the opportunity to contribute to a submission to the Australian Human Rights Commission (**AHRC**) and the World Economic Forum (**WEF**) on the White Paper *Artificial Intelligence: Governance and Leadership* (**White Paper**).

The Council acknowledges the assistance of its National Human Rights Committee, its Indigenous Legal Issues Committee, its Privacy Law Committee and Media and Communications Committee of the Business Law Section and the Law Society of New South Wales' Privacy and Data Law Committee in the preparation of this submission.

This submission supplements the Law Council's previous submission to the AHRC on Human Rights and Technology,<sup>1</sup> which commented on the impact of new technologies on human rights and the principles that should be applied for the protection of human rights amidst the increasing adoption of artificial intelligence (**AI**) and related technology. A copy of this submission is attached.

### **Does Australia need a Responsible Innovation Organisation?**

The central question posed by the White Paper is whether Australia needs an organisation to take a central role in promoting responsible innovation in AI and related technology.

The Law Council does consider that a Responsible Innovation Organisation (**RIO**) may provide value; however, a number of factors must be weighed in considering its possible establishment and it should not be considered the sole destination of responsibility or funding in this area.

The Law Council notes that emerging models are also currently being considered in other highly relevant jurisdictions. For example, in the United Kingdom careful consideration is being given to a clearing-house model.<sup>2</sup> However, no decisions have yet been made.

---

<sup>1</sup> Law Council of Australia, Submission to the Human Rights Commission, *Human Rights and Technology* (25 October 2018).

<sup>2</sup> UK Government Department for Digital, Culture, Media & Sport. *Government Response to Centre for Data Ethics and Innovation* (2019) [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/757509/Centre\\_for\\_Data\\_Ethics\\_and\\_Innovation\\_-\\_Government\\_Response\\_to\\_Consultation.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/757509/Centre_for_Data_Ethics_and_Innovation_-_Government_Response_to_Consultation.pdf).

The Law Council agrees with the White Paper's hypothesis that "Australia needs to match the rising levels of innovation in AI technologies with innovation in AI governance".<sup>3</sup>

Technological developments are advancing at an exponential rate that outpaces the ability of our current legal and regulatory frameworks to keep pace. This uneven development poses a very real threat to fundamental human rights and freedoms, particularly the right to privacy, the right to a fair trial, the right to non-discrimination and equal protection of the law and the right to liberty and security of the person. It also has the potential to disrupt labour markets by replacing human workers with automation and displace those workers unable to upskill to adapt to the new market conditions. A further risk is the increase in social inequality because of unequal access to new technologies and knowledge systems that perpetuate social exclusion.

The Law Council recognises that many of these challenges already exist. They manifest most particularly in the form of automated algorithmic-driven decision-making by government agencies and businesses. This operates in ways that may substantially and adversely affect the treatment of individuals, from rapid deployment of "internet of things" (IoT) devices and services, and through pervasive surveillance of behaviour of individuals both offline and online through use of so-called "perception AI".

AI adds to the catalogue of issues arising from "big data" and rapid and unpredictable technological innovation. There are some new issues that are particular to advanced AI. There is a new risk of opaqueness, or lack of "explainability", as to how machine learning enabled decisions are made. There is a further challenge of allocating legal responsibility for autonomous devices and liability when devices are operating autonomously (without human control or oversight) and cause harm. However, most of the issues and challenges of data driven innovations, from simpler IoT devices and mobile applications through to prospective uses of advanced AI, are highly application-specific and context-specific. The risks cannot be properly assessed, managed and mitigated without consideration of the particular circumstances of deployment and use of a particular device or application, including the level of skill and knowledge of humans involved in making decisions about how and when a device or application is used. Many of the prospective risks arise from deployment of algorithmic decision-making and pervasive data collection (whether or not also advanced AI) in specific contexts where risks of harm or unintended consequences are particularly high.

Context-specific risks require careful consideration of possible harms in particular sectors, applications and deployment environments (e.g. homes, offices, workplaces, etc.). In the view of the Law Council, it is unlikely that a generalised or economy wide responsible AI organisation would be the right body to facilitate or regulate evaluation and mitigation of context specific risk.

Moreover, there is a real risk that focus upon evaluation or and mitigation of context specific risk arising from other and more "here and now" deployments of algorithmic decision-making and pervasive data collection, whether or not classed as deployments of AI.

In addition, there is a real risk that creation and designation of a responsible AI organisation will create a false sense of security of legislatures and policy-makers, citizens and consumers. It may create a perception that "here and now" issues of context-specific

---

<sup>3</sup> Australian Human Rights Commission and World Economic Forum, *Artificial Intelligence: Governance and Leadership White Paper* (2019) 8 <[https://tech.humanrights.gov.au/sites/default/files/2019-02/AHRC\\_WEF\\_AI\\_WhitePaper2019.pdf](https://tech.humanrights.gov.au/sites/default/files/2019-02/AHRC_WEF_AI_WhitePaper2019.pdf)>.

risks and harms should and can be addressed by AI experts in a responsible innovation organisation, rather than being it being the responsibility of all regulators and organisations to adapt to an AI-enabled economy.

The Law Council suggests that, instead, government agencies, businesses, universities, regulators and civil society must be enabled to develop awareness and the capability to recognise, address, manage and mitigate risks and harms arising through data driven decision-making and technological innovation, including advances in AI. Consideration of “here and now” issues should not be deferred. Nor should current responsibilities be reallocated to address these issues, where evolution and adaptation of existing regulatory bodies, co-regulatory schemes and self-regulatory frameworks may (reliably and verifiably) mitigate risk and harm.

In summary, there is a real risk that looking at AI as a new legal paradigm, or as a fundamentally different governance challenge, might enable deferral of accountability and responsibility of current regulators and of organisations developing, deploying and using AI. Australia should be cautious in responding to calls to “do something about AI”, as though AI is an existential threat or fundamentally different innovation, rather than a new challenge requiring adaptation and evolution of existing forms of regulation to enable our society to harness benefits and manage risks of diverse forms of technological innovation.

This note of caution as to creation of a new regulatory body should not be interpreted as complacent in the face of significant new risks (as well as benefits), or as suggesting that existing regulation does not require significant adaptation. A challenge for governance of AI is how to safeguard fundamental rights and dignities, while also supporting technological development, innovation and expansion and balance the interests of various stakeholders whose interests often will not align. The WEF has noted that linear, top-down and centralised approaches to governance that were appropriate for the conditions of earlier industrial revolutions are unlikely to be applicable to the conditions of the Fourth Industrial Revolution.<sup>4</sup> The question would be – what is the most effective new form of governance? This in turn raises the question of how to stimulate development and application of bottom-up and context specific controls and safeguards that organisations developing, deploying and using AI should implement to mitigate risks and harms of deployment of AI.

### **The creation and functions of a Responsible Innovation Organisation**

The Law Council supports the establishment of a RIO subject to the cautions articulated in the previous section. The White Paper suggests that “such an organisation could combine capacity building, expert advice, governance, leading practices and innovative interventions that foster the benefits of AI while mitigating risks”.<sup>5</sup> The RIO might educate other regulators and oversee development of certification schemes, trust marks or other accreditation or endorsement of good practice frameworks, tools and methodologies that organisations developing, deploying and using AI might implement to mitigate risks and harms of deployment of AI. In this way, a RIO could promote AI innovation without abandoning supervision and oversight of technological developments in ways that are oriented to the protection of individual and collective rights and freedoms. A RIO should

---

<sup>4</sup> Klaus Schwab, World Economic Forum, ‘The Fourth Industrial Revolution: What it Means, How to Respond’ (2016) <<https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond>>.

<sup>5</sup> Australian Human Rights Commission and World Economic Forum, *Artificial Intelligence: Governance and Leadership White Paper* (2019) 8.

also be independent of both government and industry and transparent and accountable, in order to maintain the confidence and trust of citizens and other stakeholders.

The Law Council affirms its support for the adoption of human rights that are imbedded in the framework to guide the development and governance of innovations in AI technologies and therefore to guide the operation of the RIO. The Law Council considers that, with the expansion in the use of AI, it is critical that innovation be carried out in a way that is responsible, transparent and is consistent with fundamental human rights and freedoms, or is alternatively, imbedded in a human rights framework.

As laws protecting privacy have not kept pace with technological developments, consideration of the privacy aspects of AI and the protection of privacy is necessary. It is particularly important that privacy is safeguarded when using AI with sensitive information in areas like criminal justice, healthcare and in terms of facial recognition technology. Without sufficient attention to privacy concerns in the design of new technologies, there is a risk that it may not be possible to later legislate to correct harmful encroachments or redress the ensuing injustices. Any RIO should, therefore, work closely with the Office of the Australian Information Privacy Commissioner (**OAIC**).

The Law Council suggests that a key role of the RIO should be “translational”: to anticipate and articulate issues in a way that empowers policy makers, civil society and regulators to engage with issues that arise from data driven decision-making and technological innovation and do so in a way that is principles based and technologically neutral. The RIO might identify gaps in existing regulation and the current mandate of regulatory bodies and suggest how those gaps might be filled. Further, the RIO should identify where current regulation, expectations as to good practice, or regulatory sanctions are not fit for purpose.

The RIO might also make recommendations as to the appropriate scope for regulation of algorithmic decision-making and applications of autonomous AI. As noted in the Law Council’s previous submission, if AI algorithms are informed by biased or discriminatory attitudes, then “discrimination will be replicated, perpetuated and potentially even reinforced”.<sup>6</sup> This can have significant consequences for individuals belonging to particular groups.

For instance, some information that AI is designed to collect about an individual may be classified as “sensitive information”<sup>7</sup>, which is defined by section 6 of the *Privacy Act 1988* (Cth), as follows:

- (a) information or an opinion about an individual's:
  - (i) racial or ethnic origin; or
  - (ii) political opinions; or
  - (iii) membership of a political association; or
  - (iv) religious beliefs or affiliations; or
  - (v) philosophical beliefs; or

---

<sup>6</sup> European Union Agency for Fundamental Rights, *#BigData: Discrimination in Data-supported Decision-making* (FRA Focus, 2018) 10 <<https://fra.europa.eu/en/publication/2018/big-data-discrimination>>.

<sup>7</sup> The definition of sensitive information differs across jurisdictions however the Commonwealth definition is covered by section 6 of the *Privacy Act 1988* (Cth).

- (vi) membership of a professional or trade association; or
- (vii) membership of a trade union; or
- (viii) sexual orientation or practices; or
- (ix) criminal record;

that is also personal information; or

- (b) health information about an individual; or
- (c) genetic information about an individual that is not otherwise health information; or
- (d) biometric information that is to be used for the purpose of automated biometric verification or biometric identification; or
- (e) biometric templates.

These categories of sensitive information can lend themselves to biased or adversely discriminatory behaviours and decisions. As a reference point for what type of information may lend itself to algorithmic bias, the RIO might consider recommendation for adoption the above definition of “sensitive information”. The risk is that, if biased or discriminatory attitudes are built into AI, this can perpetuate bias in decision-making and potentially infringe procedural fairness or even the right to a fair hearing.

This is not to suggest that data sets containing sensitive information (for example in relation to health) should not be used in AI decision making. In the health sector, for instance, there are examples of where AI, in the form of health applications, can make a beneficial difference to an individual’s management of a health condition. Rather, the Law Council suggests that the RIO have a role in overseeing whether an appropriate balance is struck between the benefits and harms of AI.

The RIO might also develop recommendations as to socially beneficial development of AI algorithms and requirements for their transparency to mitigate against the harmful risks. Visibility into the data used in decision-making to prevent skewed data input and therefore the generation of biased data sets would be an essential feature of risk management. Data Impact Assessments may also be a means by which the impacts of AI can be assessed, so that processes for mitigation of risks are developed and implemented before data is collected and used and algorithms and AI deployed.

The RIO might also be empowered to advise as to the extent to which organisations should make available descriptions of functionality of algorithms used by those organisations where legitimate concerns about infringement of individual’s human rights may arise. While the Law Council acknowledges that there may be arguments against this from the perspective of protecting intellectual property interests or commercial interests, these considerations are to be weighed against freedom of information considerations, transparency and principles of democratic governance. The RIO would need to consider how best to undertake this balancing exercise. Algorithms may be explained without compromising any trade secret character of the underlying algorithm itself.

The RIO should have a key function of empowering civil society through education and capacity building in the development and deployment of data-driven decision-making and technological innovation. In doing so, we should preserve technological neutrality and a principles-based approach. The focus should be on outcomes not prescription.

First, attention ought to be given to educating individuals and particularly to members of vulnerable groups such as children, older people and Aboriginal and Torres Strait Islander communities.<sup>8</sup> This agenda could incorporate improving their digital ability as well as addressing the affordability and accessibility of digital tools. Secondly, it is recommended that the RIO have a role in educating AI developers to cultivate their awareness of and compliance with responsible innovation. Thirdly, the RIO's functions could include developing guidance and setting standards on the proper use of data in AI decision-making, including guidance on the ethical use of data. These educational functions would assist to protect individuals as data subjects and create general confidence in AI.

The RIO might also have a function of reviewing how algorithms are used to affect how an organisation, whether government or private sector, deals with individuals, and ensure that uses are reasonable, properly explained and tested, and relevant data appropriately protected. It may be that there also need for a regulatory body, or bodies, to be conferred with enforcement powers or and ability to exact sanctions or penalties for non-compliance with enforceable regulatory codes or stated policies of organisations. The view of the Law Council is that a generalist RIO is not an appropriate enforcement agency given the context and application specific nature of uses of algorithmic decision making and AI. Whether another regulatory body such as OAIC or the Australian Competition and Consumer Commission (**ACCC**), is conferred such powers is a matter for consideration as to good regulatory design and appropriate resourcing.

The Law Council reiterates its earlier comment that the RIO should work closely with the OAIC in relation to the privacy implications of AI governance and regulation. Similarly, careful consideration should be given to how the RIO will interact with other agencies such as the ACCC, the Australian Securities and Investments Commission and the Australian Competition and Media Authority. While there are advantages to the establishment of an independent body mandated to oversee algorithmic decision making and AI in a harmonised approach across sectors and jurisdictions, it is critical that overlaps between regulatory bodies are avoided, particularly regarding enforcement and compliance.

Please contact Dr Natasha Molt, Director of Policy, on (02) 6246 3754 or at [natasha.molt@lawcouncil.asn.au](mailto:natasha.molt@lawcouncil.asn.au) in the first instance should you require further information or clarification.

Yours sincerely



**Arthur Moses SC**  
**President**

---

<sup>8</sup> See Law Council of Australia, Submission to the Australian Human Rights Commission, *Human Rights and Technology* (25 October 2018)  [<https://www.lawcouncil.asn.au/resources/submissions/human-rights-and-technology>](https://www.lawcouncil.asn.au/resources/submissions/human-rights-and-technology).