

14 April 2020

Mr Edward Santow
Commissioner
Human Rights and Technology Project
Australian Human Rights Commission
tech@humanrights.gov.au

Dear Mr Santow,

Australian Human Rights Commission Discussion Paper on Human Rights and Technology

Telstra welcomes the opportunity to make this submission in response to the Australian Human Rights Commission (AHRC)'s Discussion Paper on Human Rights and Technology (the Paper).

Our submission reflects our thoughts on possible approaches for managing the governance of AI in Australia from a human rights perspective, including addressing what we see as the key strengths of the Paper and making some suggestions in areas where we think there could usefully be additional reflection or alternate directions considered. In summary:

- We commend the AHRC on this important contribution to the emerging discussion on how human rights can be promoted and protected amid the rise of new technologies
- We recognise the scope of the AHRC's remit is Australia, but highlight that the challenges presented by AI are global in nature. We ask that further consideration be given to how the proposals align with, or differ from, other domestic and international AI governance initiatives, as we think broad alignment would be advantageous
- We consider regulation in the AI space in Australia to be premature, and self-regulation, such as industry codes, more appropriate in the short term due to the fast-developing nature of AI and the flexible and agile regulatory response it requires
- We believe the proposal to establish a regulatory sandbox is worth exploring further. As this proposal is not without commercial risk for participants, we consider such a test space viable only if assessments are conducted as commercial in confidence
- We note with interest the proposal for an AI Safety Commissioner and suggest that the scope of this role be extended so as to act as a coordination point for all AI activity in Australia
- We recognise the importance of explainability and accountability in building public trust in AI-enabled decision-making. We do have concerns about the feasibility of the legislative and compliance instruments considered for this purpose.
- We welcome the Paper's focus on accessible technology. We support the concept of human rights by design, and the introduction of accessible technology standards, but note the need to ensure any measures taken are workable within the technology business environment.

As you are aware, the AI environment is one in which Telstra is very active across the full spectrum of applications; as a purchaser, developer, user and seller of AI-enabled systems and products, both in-house and third-party developed. As such, we have a keen interest in ensuring outcomes that reflect our organisational purpose: to build a connected future where everyone can thrive. The 'everyone' in that statement is not accidental – it is core to our purpose to ensure that all our customers can thrive in our rapidly-changing world. This means we are just as conscious of the potential impacts of AI on our most

vulnerable customers as we are of the other exciting opportunities it represents. We seek to balance the need to stimulate research, investment and commercialisation of AI in Australia – which currently lags most developed economies in this field – with the need to ensure protection of the fundamental rights to which the Paper refers.

Scope

We congratulate the AHRC on the significant stakeholder and academic research undertaken to develop the proposals put forward in the Paper. The work makes an important contribution to this emerging area of discussion in Australia and internationally. While we acknowledge the AHRC's remit is Australia, we also recognise that the challenges presented by AI are supra-national in nature and that in our globally connected world, Australians and Australian businesses will choose to use technologies developed in other jurisdictions. For this reason, further consideration needs to be given to how the proposals align with or differ from major international initiatives in AI governance.

It would be helpful to make clear, for example, how recent legislative developments in the European Union and United States compare to what is being proposed in the Paper, even where those legislative instruments do not use the same rights-based framework proposed in the Paper. While full harmonisation of regulatory regimes is beyond the scope of the Paper's work, we suggest it would be useful for any proposed Australian approach to be broadly aligned with what is happening in other developed countries, so as to facilitate technology development and allow Australia to take advantage of AI-enabled systems that are supplied on a global scale and thus minimise the cost for users to comply specifically, and potentially additionally, with an Australian framework. As a global business, we are seeking to develop a globally applicable framework for managing AI and would welcome harmonisation where possible with emerging international norms such as OECD AI Principles or the Toronto Declaration protecting the rights to equality and non-discrimination in machine learning systems.

We also recommend that Australia draw on the substantial international body of existing work in AI such as IEEE Ethically Aligned Design work or the Global Network Initiative. This could apply, for example, to the AHRC proposals relating to designing AI-informed decision-making systems to provide a reasonable explanation to individuals; or implementing a regulatory sandbox (on which we comment in more detail below). There is considerable overlap between these proposals and work already underway internationally and it would be useful for the Paper to make it clear that any Australian approach should build on, rather than duplicate, this work.

Similarly, it would be useful for the Paper to reference more explicitly the Commonwealth Government's Ethical AI framework (the Framework) and demonstrate how the AHRC's proposals support or diverge from this. The Framework is currently being piloted by a number of large Australian businesses, including Telstra, where we are investing significant resources in testing the feasibility of the Framework against a variety of real-life use-cases in our business. We are drawing on external expertise from academia, government, business and the not-for-profit sector in order to assess how we might implement this Framework or any other that is developed. While we appreciate the distinction drawn in the Paper between the law, regulation and ethical frameworks, the Government has indicated it is leaning towards using the Framework, or some modified version of it. In our view, a unified, whole-of-government approach to AI would be beneficial in providing certainty for business and consumers and we therefore suggest that the AHRC's work on AI align with what Government is already proposing where possible.

Regulation

We consider regulation to be premature in the AI space in Australia, for two main reasons: first, because we are of the view that the current pilot of the Framework (see above) will yield important insights into the feasibility and desirability of regulating the deployment of AI; and second, because we consider self-regulation, such as industry codes, more appropriate in the short term due to the fast-developing nature of AI and the flexible and agile regulatory response it requires. In our view, the dual purpose of regulating AI is to prevent and minimise harm, and to provide sufficient certainty to allow investment and technological innovation in the industry to flourish. These objectives could be achieved by agreeing on

some minimum standards that AI products deployed in Australia should meet. The development of such standards could usefully be an objective of the regulatory sandbox proposed (see below).

Consequent to our view that regulation is premature, it is our strong view that if any new regulatory framework is to be imposed, it must be sufficiently flexible as to cope with the rapidly evolving technology it seeks to regulate. For this reason, we are attracted to technology-agnostic, principles-based regulation, which promotes the right behaviours by key actors. Regulation should not become a source of competitive disadvantage, either domestically or internationally.

Naturally, any regulatory framework should be based on a best practice approach whereby a clear problem-statement, identification of options and cost/benefit analysis are undertaken to establish whether government legislation and regulation is necessary to address problems that are unlikely to be solved through commercial or industry solutions, and to whether the benefits of any such intervention outweigh the costs. As regulation has the effect of forcing businesses to allocate resources to compliance, it is critical to find the right balance between driving positive outcomes through regulation and stifling innovation by adding prohibitive compliance costs.

As we have indicated in our previous discussions with you, we consider the *Modern Slavery Act (Commonwealth) 2018 (MSA)* to be an interesting model of how a flexible, principles-based approach might work. The MSA does not impose compliance obligations on reporting entities; rather it imposes the obligation to demonstrate transparently how they are managing the risk of modern slavery in their business. The obligation is limited to large businesses that meet the \$100 million annual revenue threshold, thus avoiding deterring entrepreneurs from pursuing innovation, and applies equally to government and business. We have seen some interest from customers, since the introduction of the MSA, in greater transparency from their suppliers about how these risks are being managed. Similarly, a regulatory framework that asks businesses to demonstrate *how* they are managing the risks associated with AI, as articulated in a set of principles or ethical framework, would provide greater transparency to customers and drive positive behaviours without adding heavy compliance costs.

The Paper provides for the possibility of some ethical frameworks or regimes to be accorded a 'special legal status'. To us this seems unnecessarily complex: whether the aim is to develop some kind of trust mark, which would be advertised to consumers, or whether it is to set a *de facto* standard with which industry must comply, we consider it preferable to avoid creating a new legal status that is neither law nor regulation. We believe our proposal for a reporting regime would go a long way to addressing the question of consumer confidence in AI without imposing a legislated standard or compliance regime.

The proposal to establish a regulatory sandbox is interesting and worth exploring further but is not without commercial risk for participants. Once again, we would prefer a principles-based approach, where businesses can test emerging technologies against a series of principles, including, but not limited to, the protection of human rights. In our view, the current pilot of the Framework is an example of a sandbox for a principles-based regulation. It is essential that this work be allowed to play out and that industry feedback be considered in designing any future regulatory regime. However, a regulatory sandbox could be a useful way to develop some *de facto* Australian standards for AI-enabled products and services, subject to the technological caveats below.

In view of the significant first-mover advantage in the technology space, we consider such a test space would only be viable for AI-enabled products and services if assessments were conducted as commercial-in-confidence, with actual feedback provided to individual companies. At a later stage, a generalised description of the product or service could be used to draw broader conclusions and inform policy. We consider it unlikely that even significant financial incentives would motivate businesses to openly share commercially sensitive intellectual property in such an environment. However, if testing were commercial-in-confidence, a trust mark for successful testing could be a useful incentive for companies seeking to increase consumer trust in their products. Such a sandbox could be managed by an expanded Office of the AI Commissioner (our suggestion below under 'Governance' refers).

Governance

We note with interest the proposal for an AI Safety Commissioner, modelled on the existing e-safety Commissioner role. While we support the idea of a centralised Government agency as a reference point for AI, we suggest that the scope of such a role should extend well beyond simply AI Safety and could usefully act as a coordination point for AI activity in Australia. This could include any eventual test bed or sandbox, and the excellent proposals in the Paper related to additional funding for training, capacity building and education for professionals and government officials. It could also develop and implement the national AI education plan proposed in the Paper. In addition, such a position could interface with the AHRC on the implementation of the Accessibility recommendations in the Paper (see below for additional comments on these).

Explainability and accountability

We recognise the importance of explainability and accountability in building public trust in AI-enabled decision-making. However, in our view it may not always be feasible to provide a reasonable explanation (as defined in proposal 7) for an automated decision. We will be in a better position to comment on this notion once we have completed the Framework pilot as we are testing, during the pilot, the extent to which we are able to explain automated decisions in three real life use cases. A particular challenge arises when we use systems procured from a third-party vendor and do not have full visibility of algorithms or underlying data sets. The pilot is allowing us to work through how we might navigate this.

We also consider the proposal holds automated decision-making systems to a higher level of explainability than humans; humans are not always able to provide a reasonable explanation for their decisions and we accept this as a society, unless the outcome of the decision is to break the law. We are strongly opposed to a legislated standard of explainability for this reason and return to our earlier proposal focused on an obligation to report on how this issue is being managed, rather than compliance with a standard of explainability.

Similarly, we do not consider the absence of explainability to be a sound reason to withhold deployment of a system because it *could* infringe on a person's human rights (proposal 8). The law protects human rights. We consider it more appropriate to require businesses to mitigate the risks associated with any new product or technology they deploy, and to ensure full compliance with relevant laws, rather than placing a moratorium on the development of the technology.

The Paper asks the question whether Australian law needs to be reformed to make it easier to assess the lawfulness of an AI-informed decision-making system, by providing better access to technical information used in AI-informed decision-making systems such as algorithms. In our view, the commercially sensitive nature of proprietary algorithms makes this unworkable, and the time required between developing or iterating an algorithm and having it assessed as lawful would be prohibitive.

We are interested in further understanding proposal 10 on liability when deploying an AI-informed decision-making system. This is particularly relevant in our operating context, where we purchase third-party AI-enabled systems and deploy them in our own business, as well as selling them to customers as part of a value-added package. It remains unclear to us in some common use-cases who would be considered to be the legal person deploying the system: for example, if we give a third-party smart home digital assistant to a retail customer as part of a home internet package, are we deploying that system, or is the customer? This is where, in our view, it would be helpful to have some sort of industry code or de facto standard defining what the obligations of the operator and supplier should be. As a supplier of the system (who receives money for the end product or service) we accept that we must assume some degree of accountability. Notwithstanding this, we should be able to rely on a certain level of recourse to the supplier, and it would provide greater certainty for us and for consumers if there were rules around this, rather than it being dealt with on a case-by-case basis.

Accessibility

Telstra welcomes the Paper's focus on accessible technology and is committed to working with disability stakeholders to deliver accessible technology products to our people and customers. We support the

concept of human rights by design but note there are a range of competing interests that could make it challenging to implement this in a technology business environment, notably the rapid evolution of technology and the short time to market from initial concepts that are a significant competitive advantage in this sector. Education and awareness are critical to delivering this as a community-wide outcome and the regulatory framework should involve a combination of regulations, technology standards, and educational material for the developers, suppliers and end users of AI.

While we are encouraged by the Paper's focus on the introduction of accessible technology standards, we consider the proposal for a compliance audit to be premature. There is currently no real national standard in this space and disability advocates tell us compliance with the different standards that do exist is poor. We suggest a progressive introduction of standards, with audits conducted after the standards have been introduced, and that the technology sector must be involved in designing and testing the standards to ensure they are workable.

We would welcome the opportunity to discuss with you further any aspect of this submission.

Sincerely,

A large black rectangular redaction box covering the signature of the sender.


Executive, Government, Regional Affairs and Sustainability
Legal, Regulatory and Corporate Affairs
