March 20, 2019
By email: tech@humanrights.gov.au

Dear Commissioner,

Thank you for the opportunity to provide a submission in response to the Australian Human Rights Commission's "Artificial Intelligence: governance and leadership" white paper, as part of your Human Rights and Technology Project.

By way of background, the Digital Industry Group Inc. (DIGI) advocates for the interests of the digital industry in Australia. Its members include Google, Facebook, Twitter, Amazon and Verizon Media whose services range from search engines, content and communications platforms, and online stores. DIGI advocates for a balanced approach to regulating the online world that harnesses the tremendous social and economic opportunities digital services bring to Australia and globally, while also ensuring these services are used in a positive and beneficial way.

DIGI members recognise the importance of the issues raised in the white paper. DIGI believes that Artificial Intelligence[1] (AI) can support fairer decision-making, public safety and more inclusive and informed societies -- however, we do acknowledge its capacity for unintended consequences, and the need for solutions to mitigate against potential harm.

DIGI looks forward to further engaging with the issues raised in this white paper, and the wider Human Rights and Technology Project. Should you have any questions or wish to discuss any of the representations made in this submission further, please do not hesitate to contact me.

Best regards,

Sunita Bose
Managing Director
Digital Industry Group Inc. (DIGI)

---

[1] Definitions of "Artificial intelligence" can vary. For the purposes of this submission, we adopt the definition used by the Information Technology Industry Council (ITI) available at ITI, *Artificial Intelligence, https://www.itic.org/policy/artificial-intelligence*.

DIGI

The Immense benefits of AI

DIGI welcomes the Australian Human Rights Commission's (AHRC) acknowledgement of the benefits of Artificial Intelligence (AI) in its "Artificial Intelligence: governance and leadership" white paper. We urge the Commission to continue assessing the broader community and societal benefits of AI as it explores potential governance models, which include but are not limited to:

- **Benefits to health:** AI is helping people attain better health and well-being; a report by PwC demonstrates how AI is already transforming eight components of the healthcare system, including preventative health, diagnosis, decision-making, palliative care, research and training[2]. As one example, Google's DeepMind Health works in partnership with clinicians, researchers and patients to solve real-world healthcare problems by applying machine learning to develop software to improve clinical outcomes[3].
- **Conflict prevention:** AI is being used in emergency response and conflict prevention. The UK's Turing Institute and the US' Political Instability Task Force have been building AI capable of predicting where future conflicts may occur, drawing upon statistical data, military reports and analysing news reports[4].
- **Disability access and services:** AI is transforming inclusion and access to services for people with disabilities. AI-powered devices that use voice commands, such as Amazon Echo and Google Home are being used by people with limited sight or mobility[5], and

---

[2] PwC (June 2017) *What doctor? Why AI and robotics will define New Health.* [Online]. Available at: https://www.pwc.com/gx/en/industries/healthcare/publications/ai-robotics-new-health/transforming-healthcare.html.

[3] DeepMind (2019) *About DeepMind Health,* Available at: https://deepmind.com/applied/deepmind-health/about-deepmind-health/

[4] Weisi Guo (October 15, 2018) 'Retool AI to forecast and limit wars', *Nature: International Journal of Science.* Available at: https://www.nature.com/articles/d41586-018-07026-4

[5] The Tipping Foundation (2018) *6 ways smart home technology is benefiting people with disability.* Available at: https://www.tipping.org.au/6-ways-smart-home-technology-is-benefitting-people-with-disability/

Facebook uses AI to automatically write photo captions for the blind and visually impaired[6].

- **Information access, privacy and safety:** AI and algorithms play an important role as a sorting mechanism for the millions of terabytes of information online, enabling people to readily obtain relevant content and information. Furthermore, AI plays an increasingly important role in balancing rights to freedom of expression with online safety and security of Internet users. In fact, there are best practice recommendations, often encouraged by Australian government departments, that encourage machine learning for this specific goal. Best practice in safety-by-design and privacy-by-design include the use of behavioural and content signals can be used to identify risky users, behaviour and content, even at the point of upload or contact. For example, Twitter uses behavioural signals to identify users who target others with abuse or harassment and limits the visibility of their tweets. Similarly, a core commitment of the Global Internet Forum to Counter Terrorism, of which Facebook, Google and Twitter and members, is that digital businesses use machine learning technology to proactively detect terrorist content[7].

DIGI believes that AI can support fairer decision-making, public safety and more inclusive and informed societies, in part because algorithms and machine learning are being used by a wide diversity of private sector industries and public sector departments -- AI is by no means a technology that should only be associated with small set of highly digitised companies. However, we do acknowledge its capacity for unintended consequences, and the need for effective solutions to mitigate against potential harm. Any AI governance model should seek to maximise these benefits in a way that is compatible with advancing human rights.

Further defining the problem

DIGI believes that further analysis is needed to define the problem the AHRC is working to solve in its proposal for a Responsible Innovation Organisation (RIO), and whether there are existing bodies or efforts underway to address this problem.

We also believe that the majority of potential problems associated with AI lie in the contextual application of the technology. Measuring the real-world outcomes of these applications is a more effective way to assess their impact than examining the algorithm itself; such an approach would be more effective in mitigating harm, and would also serve to avoid problems related to the protection of valuable commercial intellectual property.

The organisation proposed by the AHRC may face a significant human resources challenges, as an extremely high level of both technological expertise in relation to AI would be required alongside highly in-depth, sector-specific knowledge of every industry and government vertical

---

[6] Matt Burgess (April 5, 2016) 'Facebook's AI now writes photo captions for blind users', *Wired UK*. Available at *https://www.wired.co.uk/article/facebook-ai-image-recognition-caption-accessibility-blind-users*
[7] Tech Against Terrorism (2017) *The Tech Against Terrorism Trustmark.* Available at: *https://www.techagainstterrorism.org/membership/trustmark/.*

where AI is applied. It is important to remember that these technologies will be applied across the economy, not only by a small number of highly digitised companies. In accordance with this breadth of application, we also believe further consultation with this wide set of public and private stakeholders is needed to inform the AHRC's final recommendations.

DIGI questions whether a centralised office would be effective in identifying or mitigating negative applications of AI, and instead we encourage further exploration of the many existing sectoral standards and regulations that are broad enough to apply to AI, or could be modernised to encompass it. For example, AI applications relating to healthcare fall within the remit of medical and health regulators, and are bound by existing rules associated with medical devices and research ethics. Similarly, existing laws relating to privacy and anti-discrimination can be useful tools to regulate against harmful applications of AI.

The white paper states "our challenge as a nation is to ensure these technologies deliver what Australians need and want, rather than what they fear"; the examples used of technology that elicit fear are largely drawn from public sector applications of AI, such as in the criminal justice system. A notable example in Australia were erroneous debt notices issued by Centrelink to vulnerable communities, based on an algorithm that analysed discrepancies between tax records with welfare payments, commonly known as "robodebt"[8]. We welcome the AHRC's acknowledgement that AI is being used by both the private and public sectors, and we believe it is critical that any proposals around oversight of this technology include any government development, procurement or application of AI.

Encouraging AI innovation in Australia

This white paper starts with the hypothesis that Australia needs to "match the rising levels of innovation in AI technologies with innovation in AI governance". It then goes on to acknowledge that i) only 9 percent of Australia's listed companies are making sustained investments in AI, lagging behind the 20 per cent in the United States, and ii) that Australia currently lags global leaders across the G20 in the adoption of automation, with 50 per cent fewer Australian firms are actively investing in automation compared to firms in comparable economies. DIGI cautions the AHRC against any regulatory approach that would stifle this important area of technological advancement in its infancy.

While we welcome the AHRC's acknowledgement that a more "dynamic innovation ecosystem" is needed, we believe that greater government involvement in the development and deployment of AI would be counterproductive in both slowing the development of such technology, and

---

[8] Shalailah Medhora (18 February, 2019) 'Over 2000 people died after receiving Centrelink robo-debt notice, figures reveal', *ABC Online.* Available at
https://www.abc.net.au/triplej/programs/hack/2030-people-have-died-after-receiving-centrelink-robodebt-notice/10821272

discouraging investment in this area in Australia. If companies are required to submit AI for external assessment to a RIO -- even if the body charged with assessing the algorithms has a dual function to promote innovation -- this will serve to discourage the deployment of AI and other digital technologies in Australia for reasons outlined below.

Firstly, the pace of experimentation and iteration within technology companies to refine their products and services is rapid, and many companies work with multiple algorithms and machine learning systems at a time; we would question the ability of a centralised body to keep up and not slow the pace of innovation.

Secondly, the finer details of how algorithms and AI work constitute highly sensitive commercial information. The prospect of having to disclose such sensitive information will serve as a deterrent to digital services, startups or any company initiating or expanding their investment in AI in Australia. This could ultimately negatively affect the variety and quality of AI-enabled products and services available to Australian consumers.

Thirdly, the compliance burden could prove counterproductive to the AHRC's aims, as the time cost and commercial concerns companies may associate with disclosing automated decision-making to an external organisation may actually serve as a deterrent to due consideration of ethical standards.

Finally, we note that AI is often developed by companies with a presence in multiple countries and utilised in different markets to where it was originally developed; we therefore encourage the AHRC to further define the jurisdictional scope of its proposals and to conduct a thorough assessment of international efforts to address these issues and how Australia can participate in or complement these efforts. There is a significant risk that Australia isolates itself from international developments and norms in this area by going it alone.

Looking to existing processes

DIGI welcomes that the AHRC has recognised the importance of "publishing professional codes of ethics for industry, drawing on pre-existing codes of practice and ethical standards in other jurisdictions and in the international sphere." We encourage the AHRC to look to existing sector-specific legislation and relevant AI ethics processes that are currently underway which may provide solutions to the problems it identifies.

For example, in last year's Federal Budget, $30 million was committed over four years to develop a national roadmap for AI, including a national ethics framework. We understand that Data61 is currently working with Standards Australia on developing standards for ethical AI use, in consultation with some DIGI members, and we would encourage the AHRC to await the outcome of this process.

We also note that the EU High-Level Expert Group (HLEG) on Artificial Intelligence is due to present a set of ethics guidelines for the development and use of artificial intelligence to the European Commission by April 2019, and the Commission will then propose how to take these guidelines forward in both EU and non-EU countries[9].

Additionally, there are important, multi-stakeholder industry initiatives to ensure ethical considerations are taken into account in the development and application of AI. For example, several DIGI members are partners of the Partnership on AI, a multi-stakeholder organization that brings together academics, researchers, civil society organisations and companies that build and use AI technology. The partnership is developing best practices in "fairness and inclusivity, explanation and transparency, security and privacy, values and ethics, collaboration between people and AI systems, interoperability of systems, and of the trustworthiness, reliability, containment, safety, and robustness of the technology."[10] The benefits of multi-stakeholder industry partnerships are already emerging; as one example, Facebook worked with researchers from the algorithmic fairness community to develop an internal tool called "Fairness Flow". It measures an algorithm's fairness across a growing number of parameters, and has been incorporated into Facebook's internal machine learning platform and is being further scaled so that the company can use it to evaluate the personal and societal implications for each product they build[11].

Such processes that relevant industries can adopt to inform their work will be an important step in encouraging the proliferation of more ethical AI. Proposing an organisation mandated with reviewing AI and machine learning before many such standards have even been introduced, let alone monitored and evaluated for effectiveness, is a leap too far too soon.

---

[9] European Commission (18 December 2018) *Have your say: European expert group seeks feedback on draft ethics guidelines for trustworthy artificial intelligence.* Available at: *https://ec.europa.eu/digital-single-market/en/news/have-your-say-european-expert-group-seeks-feedback-draft-ethics-guidelines-trustworthy*
[10] Partnership on AI (2018) *About Us.* Accessed: *https://www.partnershiponai.org/about/*
[11] Stephen Shankland (18 February, 2019) 'Facebook starts building AI with an ethical compass', *CNet.* Available at *https://www.cnet.com/news/facebook-starts-building-ai-with-an-ethical-compass/*