10 May 2024

The Senate
Select Committee on Adopting Artificial Intelligence (AI)
By Web: https://www.aph.gov.au/Parliamentary_Business/Committees/OnlineSubmission

About us

The **UNSW Allens Hub for Technology, Law and Innovation** ('UNSW Allens Hub') is an independent community of scholars based at UNSW Sydney. As a partnership between Allens and UNSW Law and Justice, the 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 UNSW Allens Hub can be found at http://www.allenshub.unsw.edu.au/.

The **UNSW Disability Innovation Institute** ('DIIU') was established as part of UNSW's strategic goal of leadership in disability inclusion. The Institute is grounded in interdisciplinary and inclusive research in which people with disability are involved as co-creators of knowledge. Uniquely among disability research centres in Australia the DIIU brings a disability studies approach to STEM as well as HASS subjects, with links to the Centre of Excellence in Automated Decision Making and Society and to the Australian Alliance for Artificial Intelligence in Healthcare.

About this Submission

We are grateful for the opportunity to make a submission in response to the Committee's Terms of Reference. Our submission reflects our views as researchers; they are not an institutional position. This submission can be made public.

We have focussed on the issues on which we have research expertise. Our main points relate to questions (b)-(f), in particular:

- the benefits of AI systems, including for disability communities, particularly where critical stakeholders are involved in co-design;
- problems associated with, but not only with, AI systems, including throughout the systems' full lifecycle;
- the need to consider problems (including those related to discrimination and disinformation) beyond a technologically scoped context, and design legislative reforms accordingly; and
- the importance of looking to other jurisdictions, but not blindly adopting flawed approaches.

b.   Risks and harms arising from the adoption of AI technologies, including bias, discrimination and error

It is now well established that the use of AI systems often inadvertently exacerbates issues of bias against population groups and communities that are already marginalised by virtue of sex, gender, class, race or other attribute, including disability. In AI-based automated systems, underlying bias and discrimination can be hidden behind the apparent objectivity and neutrality of automated social classification, evaluation and prediction. Automation bias, where it is assumed systems are objective or neutral, can make it difficult to recognise, assess and mitigate discrimination. Further, AI systems often discriminate in ways that fall outside existing protections in discrimination law.

In the case of disability, AI-based classification has been proposed to determine eligibility for disability support funding from the NDIS. However, many people with disability and their representative organisations are concerned that AI systems designed around statistical norms have difficulty with statistically anomalous populations and with the diverse, complex and nuanced realities of living with disability. Not only does this have significant implications for the provision of resources and services for people with disability, it also exerts a form of epistemic injustice by rejecting this marginalised community's own knowledge of their own bodies and needs. Analogous forms of representational and epistemic injustice are likely to emerge for other marginalised groups. Where new systems are introduced in areas such as disability support, it is important to do so cautiously and to evaluate their use in partnership with affected communities, monitoring for harmful impacts including over the longer-term.

However, it is important to recognise when designing any response to concerns about bias, discrimination, and error, that these problems are not associated solely, or even primarily, with a particular technology such as AI. Indeed, looking at definitions of AI such as those in Europe's *AI Act*, the targeting of the response is poor. Bias and discrimination are frequently the result of data-driven inferencing more broadly, particularly in contexts where data quality is low or poorly aligned to the context of its use. Bias and consequential discrimination can result from AI, but also from more manual statistical methods. Indeed, they can result from human decision-making, the main difference being that human discrimination is better managed through legislation. What is needed is not a response to AI or any particular technology, but rather a rewriting of discrimination law to better manage the broad range of contexts in which decisions are discriminatory, whether those decisions are human, machine or (increasingly) a combination of both.

A similar point applies to error, although that is a much broader problem (indeed bias and discrimination might be described as a kind of error). For certain decisions and processes the impact of error is high, including in contexts where safety is critical. In some contexts, the use of AI increases the risk of error while in others it reduces the risk of error. Focussing on error only in the context of AI is thus not a helpful starting-point and could create a regulatory skew that increases overall error. Instead, it would be more helpful to focus on context and regulate with the goal of error reduction in critical contexts in a relatively technology-neutral way.

It should also be borne in mind that bias and discrimination are not only an issue in terms of the output of AI (and other) systems but also an aspect of production. In particular, one can ask about the unfair impact of the material (and human) infrastructures that bolster the use of AI technologies across the world. While the discourse on AI frequently imagines fewer human interactions with the growth of the technology, insights from on-ground experiences show that a large invisible human labour force is involved in the processes of data generation, data preparation or labelling, and data verification. This kind of invisible human labour, particularly in the production and proliferation of AI technology, is often ignored and is steeped in precarity. As such, there is a need for reflexive, long-term policies that consider the material conditions of production of AI as for other products and services. This should consider the problems of low remuneration, environmental risks, exploitative labour practices, and the general lack of welfare measures. Studies conducted in and around the Covid-19 pandemic also demonstrated the impact of mass adoption of AI (including automation) technologies on supply chain logistics, including mass layoffs and reorganization of logistics to relegate most labour to the 'lowest value denominator' in supply chains (like gig economy workers). Studies done on countries like India, which are fast emerging as hubs for the global data annotation industry (the market serviced by India is estimated to be about $7 billion by 2030 with a workforce of 1 million people), have consistently revealed unfair and exploitative working conditions and inequitable contracts particularly affecting women and people from marginalised caste and class backgrounds. Mass protests across several countries including India, Brazil, Mexico, Costa Rica, Peru, Chile and Ecuador, against global companies that are reliant on AI technologies have revealed a familiar pattern of demands for higher wages, better working conditions, protective equipment at workplace, and improved contractual terms.

Artificial intelligence, despite its name, still involves people. What technological change in this field as in others generates is reorganization of existing social and material conditions of labour which can disadvantage already precarious and vulnerable worker groups. As such, there is a need for more transparent and context-driven studies to understand the full extent of human participation, and the material conditions that govern the use and adoption of AI technologies. Given that AI technologies developed in wealthier countries, including the U.S., Europe, and Australia, are reliant on the labour force in poorer countries, it is particularly important that policies framed in Australia consider the material conditions of AI technologies from across the world.

c.   Emerging international approaches to mitigating AI risks.

At present, the most well-known and advanced regulation of AI is likely to be found in the European Union, *the Artificial Intelligence Act*. This Act defines "artificial intelligence" and takes a risk-based approach to regulating it, with varying obligations on providers and users depending on how risks are classified. While the EU AI Act undoubtedly has merits, it also suffers from the "define the technology and regulate it" approach that is based on a moment-in-time conception of artificial intelligence.

Australia could instead look to some of the thinking in Switzerland, which encourages a distinct approach to that operating in the EU. In particular, a Swiss position paper argues:

The challenges posed by algorithmic systems are manifold and often have a new dimension or quality, but they are not unique to such systems. Therefore, these challenges should not be covered by a general "AI law" or an "algorithm law". Instead, a combination of general and sector-specific standards is appropriate. The focus here is on the selective adaptation of existing laws.

The US approach is still emerging as the relevant Executive Order requires regulatory action from different agencies. Like the EU approach, however, the focus seems to be on the technology rather than broader problems such as bias and error.

The "Global South", while not presenting a uniform approach, takes a [different approach](#) to AI regulation, predicated by the inclination to harvest AI potentialities to provide mass employment and leverage it for widespread use throughout the economy. AI regulations therefore tend to develop to support the industry, rather than be focussed solely on harms. Countries like India adopt a techno-legal form of regulation that seeks to govern its digital public infrastructure through a combination of different laws including the nascent data protection law, competition regulation and consumer protection, with an interplay of different sectoral laws. Policymakers view technology as a critical tool for achieving economic and development priorities and a strategy to assert the country's leadership role on the global stage. This [discourse](#) is also dominated by 'sovereign AI' conversations, that push the agenda of self-reliance, regulation supporting domestic industries, framing of data as a national asset and hence data localisation, and nationalist undertones in AI/technology regulation not completely divorced from electoral politics.

d.   Opportunities to adopt AI in ways that benefit citizens, the environment and/or economic growth, for example in health and climate management

The use of AI has enormous potential benefits for people with disability. There is considerable interest in the use of AI-based systems to enhance existing assistive technologies or offer new modes of accessibility (see, for example, the rise of automated captioning or audio description for people with hearing or visual impairments). Algorithm-based systems also, in principle, have the potential to improve resource allocation by making processes more efficient and comprehensive, thereby ensuring a fairer distribution of support. At the same time, as discussed earlier, they also carry the risk of exacerbating existing, or at worst introducing new, inequities. Beyond the focus on specialised assistive technology, it will be important to ensure that general AI platforms remain accessible to people with disability.

The problematic assumption that technology, including AI, can 'solve the problem' of disability inclusion is widespread. In some cases, a novel assistive technology may improve accessibility for some while raising new issues for others. For example, for people with visual difficulties, voice activated technologies can improve accessibility to certain technologies and platforms, but at the same time make them inaccessible to D/deaf or hearing-impaired people. An over-reliance on sophisticated technology, that is almost always individualised and expensive, can increase rather than decrease the exclusion of poorer people with disability, who might otherwise benefit more from more conventional, cheaper, and often collective solutions.

As AI systems are introduced, there needs to be a structured process aimed at increasing public understanding of and engagement with these tools as well as digital systems more generally. This process needs to be conscious of the interests and needs of people with different kinds of disability and impairment, including people with intellectual disability. This again speaks to the need for the integration of co-production and co-design with disabled people when designing, developing, and implementing new technology.

e.   Opportunities to foster a responsible AI industry in Australia

The best mechanism for fostering a responsible AI industry in Australia would involve a combination of:

- revamping legislative protections against discrimination so that they mandate and encourage processes for non-discriminatory decision-making in critical contexts;

- enhancing consumer protections, particularly against targeted manipulation;

- prioritising privacy law reform;

- using administrative law reform, assurance frameworks, technical standards, and government procurement protocols to leverage the government's purchasing power in favour of responsible AI;

- educating future Australians to better navigate a world of data-driven inferencing (about them and content personalised for them), including how AI tools such as LLMs work and their limitations;

- viewing responsibility in a broader context that takes into account the entire AI lifecycle;

- working with impacted communities, including disability communities, as critical stakeholders;

- supporting Australian involvement in international standard-setting for responsible AI. Standards Australia's Committee for AI standards (IT-043) has been influential internationally including in the development of international standards for 'trustworthy' AI (which is the term used by ISO/IEC JTC 1 SC 42 WG 3). Australians are also involved in IEEE standards development. The government should support such involvement, providing financial support for civil society organisations to participate in committee work.

We dealt with this issue at greater length in our submission to the Department of Industry, Science and Resources on safe and responsible AI.

f.   Potential threats to democracy and trust in institutions from generative AI

This problem has, over time, been tacked through different lenses, including foreign interference, digital platforms and social media. The challenge democracy faces comes from all of these, as well as from advances in generative AI. But looking at the issues separately (from the standpoint of different committees) obscures the broader problem and makes it more difficult to solve.

Nevertheless, there are a range of issues to consider associated with generative AI in particular, including:

- the problems associated with deep-fake images and videos that change perceptions of truth;

- the increased speed with which lies can travel online when text is automatically generated and tailored for different audiences, sometimes by powerful foreign actors; and

- the operation of freedom of information laws in a context where information held by government is diffuse, stored through complex multi-dimensional word associations in large language models rather than in "records" that citizens can access.

In this context, the World Economic Forum's 'Global Risk Report 2024' highlights how the widespread impact of misinformation and disinformation using technological tools might have a delegitimising impact on governments, given major elections in several countries across the world, including the UK, US, India, Pakistan, Bangladesh, Mexico, and Indonesia between 2024 and 2026. The use of AI, and in particular generative AI is anticipated to be crucial, with impacts going beyond the elections - polarising public discourse from issues relating to public health and social justice, undermining truth, and laying the groundwork for more censorship.

India, marked as the riskiest country in the world in the immediate context of its ongoing elections, with 1.4 billion affected people (with nearly 50% internet penetration), presents an important case study. Following the massive proliferation of fake news propagated by AI technologies during the Covid-19 pandemic, the Indian government has responded with more stringent internet and new media (including intermediary) regulations. But this has failed to prevent deepfakes being used by and against political parties, the building of counterfeit websites, and even predicted stock market manipulations. In the last few years, India has seen a 214% rise in fake news (earning the moniker of the 'disinformation capital of the world'), triggered by an increasing use of social media, internet penetration, and the ready availability of AI technologies. This has triggered polarisation along already fraught ethno-religious lines, communal riots and pogroms, lynchings and systematic targeting of vulnerable population groups, and weaponisation of fake news and disinformation by electoral politics and political parties. Subsequently, the response from the State was marked by more centralisation of regulatory power, an increase in executive power, promulgation of emergency power provisions and increased censorship on digital media companies (including independent media outlets and journalists) and intermediaries. Some have even argued that India's regulatory response to online misinformation violates international human rights law. Rampant proliferation of misinformation and disinformation made increasingly complex and efficient using generative AI tools ushers more censorship laws and ill-considered legal regulations to counter the harms swiftly.

Where the issue is framed around technology, such as AI, it is often suggested that a government response needs to be swift and focus on the technology of concern. In the context of misinformation and disinformation, this would be unwise. There is a need to

move slowly, considering varied case studies from across the world before deciding on Australia's legal and regulatory response. There is also a need to avoid focussing only or primarily on one technological aspect of what is a broader problem. Technology will continue to evolve and will continue to impact the information ecosystem – some of this will be classifiable as AI (under some definitions) while others will result from different techniques. It is thus important to focus on the challenge broadly rather than just those aspects of it associated with AI.

Yours sincerely,

Lyria Bennett Moses, Jackie Leach Scully, Shohini Sengupta, Katharine Kemp