Submission No 16

INQUIRY INTO ARTIFICIAL INTELLIGENCE (AI) IN NEW SOUTH WALES

Name: Mr Malin Kankanamge

Date Received: 16 October 2023

Intro

Thank you for the opportunity to make a submission and share my views about how AI might impact NSW, including the risks and challenges it presents.

As a member of the NSW community and someone working within the financial technology sector in Australia, I've been thinking a lot about AI and how it is already affecting my work and my community, and the impacts it could have in the future.

When I read NSW's AI policy and assurance framework, I appreciated that the NSW government has been ahead of other jurisdictions in understanding the pace of change of AI and its transformative nature. I want the government to make the right calls in a timely way, and stay ahead of emerging issues and risks from AI. The recommendations I provide below hope to contribute to that.

Elections and misinformation

Further to terms of reference 1.(g) - I think advanced AI capabilities are likely to pose a significant threat to the ability of NSW and Australia to maintain strong and fair democratic institutions.

The fact that generative AI materials (like images created by Midjourney, Stable Diffusion, dall-e, etc and up and coming generative video and audio models which can recreate human voices perfectly in many languages) are already being used, without transparency, to influence Australian elections is deeply concerning. Recent news reports that AI-generated images, seemingly of Aboriginal and Torres Strait Islander people, were being used to advocate against the Voice to Parliament shows this isn't just a speculative concern. These tools are increasingly capable of creating perfect replicas of people's faces, bodies, voices and even movements such that it is impossible to distinguish between a real person and a generated model of a person doing whatever has been prompted - famously called "deepfakes" online.

AI capabilities are improving at an exponential pace and without serious intervention and regulation these models will become incredibly powerful tools for influence and deception. **Without interventions, it is inevitable that AI will become ever more involved in elections and political campaigns**. Without safeguards in place, malicious agents could use these tools to deceive a large enough part of the population to change the outcome of an election.

Although election interference has always been possible, AI will let more people do it on a larger scale. Urgent action is needed to address election integrity. Our democracy is the most valuable thing we have, and AI is a growing risk that demands action.

Meaningful action to protect elections needs to target AI developers and AI deployers, not just the specific conduct involved. **For other things that can be misused - from guns to cars - we ban dangerous conduct by individuals at the same time as regulating the technology itself**. In some cases that regulation effectively bans the technology (for instance, automatic weapons are limited to law enforcement and military applications) or allows it only if it meets strict safety standards (for instance, cars have requirements about protecting their occupants and other road users). AI should be treated in the same way.

Specifically, I think NSW should:

1. Prohibit the use of AI-generated content in relation to elections.
2. Require in general that all AI-generated content is "watermarked" so that it is transparent to viewers where content – particularly

pictures and videos – is generated by AI.
3. Require developers and deployers to only make AI available in NSW if it meets safety standards – including watermarking.
4. Impose strict punishments on developers and deployers who make AIs available in NSW that don't meet safety standards.

I think this kind of approach would help maintain election integrity in the immediate term, and I think it would be a significant step towards dealing with the dangers of AI in other contexts. **In the same way that NSW sets rules for the road or construction or retail, it needs to become normal that we set rules for AI.**

Banning or regulating high risk AI

Terms of reference 1.(k) asked after measures other jurisdictions, both international and domestic, are adopting in regard to the adaptation to and regulation of AI.

One of the global developments that is most promising is the creation of "national laboratories" to enable technical tests on AI models, provide technical reports and provide ongoing monitoring and assurance. Singapore has established the AI Verify Foundation, the EU has created a Centre for Algorithmic Transparency, the UK has a Foundation Model Taskforce and the Tony Blair Institute for Global Change has proposed that the UK create "Sentinel" with a similar goal.

Without a similar lab in Australia or in the region, deploying trusted and safe AI in Australia might become impossible as capability and capacity increases.

NSW has an enormous pool of very talented and capable AI Engineers and researchers at public and private institutions across the state and is well positioned to collaborate with other jurisdictions to create or support a national laboratory for AI safety, modelled on international best practices. This will allow our researchers to stay at the forefront of AI safety developments and fund necessary research on how we can steward AI models towards the best interests of Australians. Additionally, a national laboratory in NSW is likely to attract some of the best and brightest engineering and research talent from other countries in the Asia-Pacific region which will be a great boon for NSW and Australia at large.

This approach is exciting because there's international best practice to follow, and **NSW could use the laboratory to ensure the AI products it uses are safe and can be subject to effective ongoing monitoring and assurance.**

Perhaps even more importantly, we are already seeing various kinds of dangerous and risky AIs. If we had a trusted national laboratory, it could assess AI products before they go to market. In the same way we don't let cars on our roads without them going through safety tests, a lab like this would allow us to block AIs until they've passed appropriate safety tests.

Conclusion

Overall, we know that we're on track for AI technology that continues to accelerate and transform our society. I hope that NSW continues to frequently re-think how best to configure its AI policies, adapt to emerging evidence, and encourage the other governments of Australia to do the same. We don't know today if AI is trending to make things very good, or very bad. What we do know is that we need vigilant governments that are watching these trends and are ready to act.