

**Submission
No 8**

INQUIRY INTO ARTIFICIAL INTELLIGENCE (AI) IN NEW SOUTH WALES

Name: Mr Laeeque Jamdar

Date Received: 5 October 2023

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 younger member of the NSW community, 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. I am worried about the state of the world when I have children.

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 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.

Our democracy is the most valuable thing we have, and AI-produced misinformation and disinformation puts it in jeopardy. I understand that Australian governments are trying to "strike a balance" between safety and prescriptive regulation and that governments want to take a "technologically neutral" approach. However, when the thing we value most is at risk, we should have a safety-first approach.

We've already seen reports of disinformation and deep fakes being used to misinform the public about events and influence elections. And those same misinformation tools are being used to contribute to fraud and scams and even manipulate shareholders.

The NSW Parliament should act decisively to protect democracy. Democracy is too valuable to risk for vague promises of economic benefit. We can always unwind regulations once we have the tools that make AI safe.

This might look like:

- Prohibiting AI-generated content relating to elections.
- Requiring all businesses operating in NSW to be transparent when using AI-generated content.
- Requiring all AI-generated audio, pictures and videos to be "watermarked", so transparency requirements can be enforced.
- In general, having minimum safety standards for AIs operating in NSW, and requiring AI developers and deployers to meet those standards. A regulatory framework that sets minimum safety standards could be used now to protect elections and address other harms we know about. Building a regulatory framework would also let us move quickly as new dangers emerge.
- Punish developers and deployers who make AIs available in NSW that don't meet safety standards.

I understand that one objection to this might be that developers don't know how to make AIs that can't be used to subvert democracy (or avoid other "misuse" and "dual-use" concerns). If an AI can generate information, perhaps it's hard to stop it from generating disinformation. I don't think we should be persuaded by this argument. If AI can't be reliably safe, it's not ready to be let loose in our society. Only by setting clear expectations for developers and deployers can we encourage them to invest in AI safety and to prioritise addressing safety concerns. Once AI is safe, we can pursue its benefits. Trying to pursue its benefits before we know it is safe doesn't

make sense.

NSW's AI assurance framework (page 23) includes consideration of possible harms of an AI system, and frames that in terms of the residual consequence after mitigations are applied. NSW should be commended for going further and considering secondary or cumulative harms (page 24). This is often neglected, and second-order effects can often be much more significant than primary effects.

That said, there is room for NSW to improve what it considers a secondary harm. Specifically, the bulk of the future risk of AI systems could turn on the values and priorities of the developers of frontier models - including factors like how committed they are to safe and ethical AI systems and how much they are investing in AI safety research. A future where commercial incentives encourage AI developers to set aside safety considerations is a much worse future than one in which they prioritise it.

Phrased another way, if NSW signs contracts with AI developers who do not take AI safety seriously, the secondary harm could be very significant. In light of that, NSW should update its guidance regarding secondary harms to include the implications of engaging any particular AI developer – including the reputational benefit for that developer and the implications of it receiving further funding. If NSW is reaching an agreement with a frontier model developer, its main considerations should relate to the demonstrated commitment of that developer to long-term AI safety. NSW should only deal with AI developers with strong commitments to AI ethics and AI safety – including demonstrated investments in and commitments of computing resources to longer-term AI safety considerations.

Finally, I understand that there is a range of views about AI. Some people think that it could be an existential risk, and others think that it will solve all our problems. What I think everyone agrees on is that we need more investment in AI safety research. If existential risk is possible, AI safety research could save humans from extinction. Even if existential risk is not seen as a priority, AI safety research could avoid present harms and ensure AI is able to understand our intents and operate efficiently and effectively. I believe that it's the responsibility of every government globally to address new technological risks, and steps taken by NSW can motivate other states and countries to take steps as well. In that context, anything NSW can do to foster research into those issues is going to make a positive difference.

References:

This 2020 NSW paper about regulation talks about 'technological neutrality and striking a balance between regulation and safety. [Apo-nid307176.pdf](#)

Europol, 28 April 2022, [Facing reality?](#) Law enforcement and the challenge of deep fakes.

[An A.I.-Generated Picture Stokes a Stock Market Plunge - The New York Times \(nytimes.com\)](#)