**Submission by Joseph Tan**

**Date: 25 July 2023**

**1: Definitions**

Do you agree with the definitions in this discussion paper? If not, what definitions do you prefer and why?

Yes, I agree.

**2: Potential gaps in approaches**

What potential risks from AI are not covered by Australia’s existing regulatory approaches? Do you have suggestions for possible regulatory action to mitigate these risks?

I see AI/ML has societal-changing technologies and there is a risk that while we can attempt to regulate the technology in Australia, the nature of global commerce and communication will mean that foreign AI system that do not conform to Australian Ethical principles may become more prevalent. We may want to regulate these kinds of products/services or systems, but visibility of issues in large and diverse AI systems may be very difficult to uncover.

Are there any further non-regulatory initiatives the Australian Government could implement to support responsible AI practices in Australia? Please describe these and their benefits or impacts.

Education I think would be particularly important. We want the general public (particularly young people who will drive further innovation and different applications of this technology) to understand how these systems function, how to manage the risks and to fundamentally understand where the strengths of AI/ML technology are, and how it differs to human capabilities. This area of development will likely progress rapidly.

**4: Target areas**

Given the importance of transparency across the AI lifecycle, please share your thoughts on

* where and when transparency will be most critical and valuable to mitigate potential AI risks and to improve public trust and confidence in AI?
* mandating transparency requirements across the private and public sectors, including how these requirements could be implemented.

Do you have suggestions for:

* whether any high-risk AI applications or technologies should be banned completely?
* criteria or requirements to identify AI applications or technologies that should be banned, and in which contexts?

What initiatives or government action can increase public trust in AI deployment to encourage more people to use AI?

Conduct education campaign as described in response to (2).

**6: Risk-based approaches**

Do you support a risk-based approach for addressing potential AI risks? If not, is there a better approach?

Yes. This would be similar to many other industries (nuclear safety, industrial plants etc).

What elements should be in a risk-based approach for addressing potential AI risks? Do you support the elements presented in Attachment C?

Attachment C is found on page 40 of the discussion paper.

The attachment provides guidelines for human-in-the-loop and where it may not be appropriate. One additional point I wish to express is that there is a possibility of the introduction of general AI systems with at least the capability of high-functioning humans. Fundamentally there is no reason why this would not be possible and it will be a matter of when, not if, that it will occur. We would need to ask ourselves philosophically and as a culture, what is the role of people in a world where our machines will likely be more capable in many areas?

There is a risk that our society will be unprepared for the societal and psychological impact of machines that are similar to conscious entities if this were to become reality. To mitigate this, we need to prepare clear and consistent guidelines (much of it is discussed in the paper) to ensure responsible change, however the technology develops.