**Response to *Supporting Responsible AI* discussion paper**

**Background**: Dr Andrew C. Cullen, Postdoctoral Research Fellow in Artificial Intelligence Security at the University of Melbourne. I have been working in this field for four years, and I have experience in both academic and industrial applications of AI.

**Response Summary:**

This call for submissions is almost certainly going to lead to a significant number of comments regarding the risks of strong, or superintelligent AI. As an expert working in this field, I can say with great confidence that that the risks with AI are not from strong AI, but rather from weak AI - of the kind that is already being used now in industry - being applied in ways that don’t reflect the values and interests of the Australian public. Concerningly in Australia we simply don’t have a broad enough base of expertise for practitioners of AI to properly ask questions of how their systems are being deployed,

The secondary concern that I would like to emphasise is that without keeping up in the global arms race for computing power (of the top 100 supercomputers in the world, only 2 are in Australia [<https://www.top500.org/lists/top500/list/2023/06/>] and we are consistently falling down the list of supercomputing capacity) and AI talent, Australia will loose the ability to control the values of AI systems deployed within our country. A world where global AI is built according to American or Chinese interests and data is one that will lead to fundamental and deleterious changes in things that matter to Australians, and will make it significantly harder for Australian values and interests to be preserved long term. This is why Europe is investing so heavily in the idea of European AI - because they wish to preserve European values within the AI systems that their citizens will interact with. This is a paradigm that Australia must consider into the future.

**Potential gaps in approaches (Q2-4)**

**2. 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?**

The risk that the values of Australian businesses, and the Australian public, do not align culturally with AI systems developed in China, or by American businesses.

**3. 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.**

There is no other answer to this besides money, to be honest. Australia needs a pipeline of talent that does not result in talented Australian experts leaving the country for better opportunities (which is a significant factor at the moment), and we need an ability to better develop younger talent in the field.

All of this though is immaterial if this talent pipeline isn’t supported by computing infrastructure, which is a core part of nurturing, developing, and cultivating talent within this space.

**Responses suitable for Australia (Q5)**

**5. Are there any governance measures being taken or considered by other countries (including any not discussed in this paper) that are relevant, adaptable and desirable for Australia?**

The concept of European AI, in that AI systems should be developed in alignment with the cultural, financial, and linguistic values of the European union bloc, is a concept that Australia should be broadly echoing. Without taking such a position, our cultural values will not be reflected within the systems that are rapidly becoming a core and inescapable part of our everyday lives.

**Target areas (Q6-11)**

**7. How can the Australian Government further support responsible AI practices in its own Agencies?**

Government needs to interact more broadly with academia to help develop talent and expertise in responsible AI. At the moment, as far as I’m aware, Defence is the primary driver of such interactions, and even that is thin on the ground, and potentially drying up with time.

**10. Do you have suggestions for:  
 a. Whether any high-risk AI applications or technologies should be banned completely?  
 b. Criteria or requirements to identify AI applications or technologies that should be banned, and in which contexts?**

The genie is out of the bottle here, and to some extents a ban is impossible. However we need to work out how to build safe AI systems, and how to promote their use by business and government.

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

If Australia wants to be a country in which there is public trust of AI, we should be positioning ourselves to be leaders in responsible, ethical, Australian-values driven AI. At the moment we’re not anywhere close to that, and such a positioning will not happen organically. It will require work, expertise, and time - and at the moment we’re rapidly running out of time.

### **20. Should a risk-based approach for responsible AI be a voluntary or self-regulation tool or be mandated through regulation? And should it apply to: a. public or private organisations or both? b. developers or deployers or both?**

I fail to see how voluntary or self-regulatory schemes will work. AI systems are often systems that are incredibly difficult to scruitinise, and companies and individuals will *not put the effort in to fully understand the responsibility of their AI systems unless they are forced to*. A voluntary scheme, or a self-regulatory scheme will simply lead to people doing the bare minimum, which is not enough to ensure that AI systems are behaving ethically, and in a fashion that reflects the values and interests of the Australian people.