The XXX requested that its XXX provide feedback on the Safe and responsible AI in Australia document.

The following represents a synthesis of this material.

Definitions

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

   While the three examples shown are good, it does not empathise sufficiently the ubiquitous nature of AI.

Potential gaps in approaches

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?

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.

4. Do you have suggestions on coordination of AI governance across government? Please outline the goals that any coordination mechanisms could achieve and how they could influence the development and uptake of AI in Australia.

   AI Governance processes should reflect existing systems and processes as AI is just one tool of many that should sit within broader governance processes.

   This would include ISO Standards i.e. ISO 38507, titled "Governance implications of the use of artificial intelligence by organizations," which provides guidance for members of the governing body of an organization to enable and govern the use of AI but can then also sit within an organisations broader ISO processes.

   This then serves to inform regulators, certifiers and auditors.

   This provides a framework for government and non government organisations to manage use of AI but not the technology.

Responses suitable for Australia

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?

Target areas

6. Should different approaches apply to public and private sector use of AI technologies? If so, how should the approaches differ?

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

   Development of broad AI literacy
   Upskilling internal public sector technical support.
   Not trying to do everything in house – Government employees cant keep up with the depth and breadth required at the pace required especially in niche/ expert fields.

8. In what circumstances are generic solutions to the risks of AI most valuable? And in what circumstances are technology-specific solutions better? Please provide some examples.

9. Given the importance of transparency across the AI lifecycle, please share your thoughts on:
   a. where and when transparency will be most critical and valuable to mitigate potential AI risks and to improve public trust and confidence in AI?
   b. mandating transparency requirements across the private and public sectors, including how these requirements could be implemented.

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?

11. What initiatives or government action can increase public trust in AI deployment to encourage more people to use AI?
    - Increase hands literacy development by individuals and groups
    - Provide hands on opportunities to engage with AI in simple language in practical and meaningful ways e.g. hackathons, SME pitches to Business, Board and Executive Programs etc

Implications and infrastructure

12. How would banning high-risk activities (like social scoring or facial recognition technology in certain circumstances) impact Australia's tech sector and our trade and exports with other countries?

13. What changes (if any) to Australian conformity infrastructure might be required to support assurance processes to mitigate against potential AI risks?

Risk-based approaches

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

    Yes but:

As is common with AI, people tend to look at it through completely new lenses, rather than use existing frameworks and fine tune them. We use tools everyday and risk manage their use – human intervention, decision support etc

We are teaching companies to responsibly use AI, the same as a learner driver.

Some drivers have been in go karts from 5 (or driving around the farm), used to regulations, now just have a new beast and thus a few new regulations (so not much to do).

Others are completely new to driving motorised vehicles. So there are base regulations, simulations, then L and P plate times (which may involve driving lessons). Regulations are tighter at this stage, and they can lose their license very quickly. Then they pass the driving test, and are carefully monitored, with guard rails in place. Monitoring is a combination of fixed and random monitoring, with opportunities to lose the license. We also have many signs to remind people of what they need to adhere to. So all that needs to be in place.

We also need to account for the people who ignore the 'road ends' sign, continue on the dirt road, drive around the fence at the end (yes we need the fence), and drive of the cliff. So we need the fence, and the ambulance, as well as monitoring on the dirt road section. Cars also are monitored for defects (more in some countries than others), particularly as they get older.

Most of the suggested parts of the plan are about building the road rules and vehicle safety standards, with a little on L's and P's.

Almost none of it is monitoring in production, guardrails around this, and escalating steps of enforcement (nor any ambulances).

Critically there is no 'monitoring as the algorithm ages', which is akin to 'phase 4 research or post marketing surveillance' step.

Given the data is needed to run the algorithm, part of the guardrails is that it is continuously monitored and reported, with expectations of improving functionality over time (rather than static), as both the data size/quality, and algorithms should keep improving.

For any algorithm to come to production, the company involves need to submit its testing and maintenance/improvement strategies, and report against these (with yearly review of both KPIs and reporting, with the intention both will be upgraded every year as we all learn).

In summary the document is reasonably good at the 'build the car and roads, and start driving on the Ps' part' but the additional components after that are missing.

It also misses the fact that these AI products will constantly change in a production environment (unlike cars/machines/drugs), and thus the 'monitoring in production' phase will be far more intense, nuanced, constantly evolving and critical.

15. What do you see as the main benefits or limitations of a risk-based approach? How can any limitations be overcome?

    Risk based approach relies on user Business Governance.

    Many tools are direct to user vs to business. The draft tool identifies that the user must be trained. Practically this is problematic – who trains the user and how in a direct to market online tool? More realistically – tools should be available to enable self-directed online tool use.

    Monitoring and documentation needs to define who – The end user/ the using company/ the developer. The framework should clearly define who each element is targeted at and the outcome of the risk mitigation strategy.

16. Is a risk-based approach better suited to some sectors, AI applications or organisations than others based on organisation size, AI maturity and resources?

    Given the international nature of AI development and deployment and its ubiquitous breadth the challenge will be to be able to quickly identify where AI is low risk and minimise the administrative burden. Too much administrative burden will create two scenarios a) lack of compliance with administrative tasks, lack of local product due to over regulation.

    Where market segments try to “protect” the vulnerable with regulation, the opposite may occur where the vulnerable are denied access due to companies failure to comply or the cost for the vulnerable will rise to cover administration, with cheaper access for the masses.

    There are areas where the regulation should be incorporated into existing governance processes based on risk e.g. therapeutic goods and defense.

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

18. How can an AI risk-based approach be incorporated into existing assessment frameworks (like privacy) or risk management processes to streamline and reduce potential duplication?

19. How might a risk-based approach apply to general purpose AI systems, such as large language models (LLMs) or multimodal foundation models (MFMs)?

    While not directly referenced part of the risk for generalised public consumption models should sit within the existing frameworks of informed consent, expressed in “simple terminology”.

20. Should a risk-based approach for responsible AI be a voluntary or self-regulation tool or be mandated through regulation?
    It should depend of the level of risk

    And should it apply to:
    a. a. public or private organisations or **both**?
    b. b. developers or deployers or **both**?

    and
    c. C. ***the individual consumer***