Supporting responsible AI: discussion paper
Submission from Gadens

4 August 2023

Introduction

Gadens is pleased to have the opportunity to make a submission in response to the questions contained in the 'Safe and responsible AI in Australia' discussion paper published 1 June 2023 by the Department of Industry, Science and Resources (**Discussion Paper**).[1]

Our submission responds to and includes our comments and observations regarding some of the proposals in the Discussion Paper which we consider would have the most significant impact on individuals, our clients and business and/or will raise important commercial or legal compliance issues.

Any defined terms used in our submission have the meaning given in our submission and are otherwise defined in the Discussion Paper.

We consent to this submission being published on the Department of Industry, Science and Resources' website.

Definitions

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

We consider that in the context of definitions, the main focus should be on finding a single future-proof and technology neutral definition of AI, which can be supplemented by references to the AI techniques and approaches outlined in the Discussion Paper, such as large language models and automated decision-making (e.g., in the form of annexures or rules to potential AI legislation which would allow for flexibility and adaptability for both voluntary and regulatory approaches). Given the changing nature of AI technologies and underlying applications, we believe that a less prescriptive approach is preferable over a fixed set of definitions. In relation to the proposed definition of AI we further submit the following alternative definition:

Definitions in Discussion Paper	Gadens proposal
Artificial intelligence (AI) refers to an engineered system that generates predictive outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives or parameters without explicit programming. AI systems are designed to operate with varying levels of automation.	AI refers to an engineered system developed with one or more of the approaches listed/defined in Annexure X/Rules X that generates outputs such as content, forecasts, recommendations, or decisions influencing the environment they interact with.

[1] Department of Industry, Science and Resources, 'Safe and responsible AI in Australia' (Discussion Paper, 1 June 2023) 34-5.

The proposed definition departs from the definition of AI proposed in the Discussion Paper in several ways:

1.1 Element of predictability of outputs to be removed

We consider that the element of predictability is too restrictive and, to a degree, too subjective. While predictability is an initial design feature of many AI applications, it is questionable whether this element can be sustained e.g., in the context of generative AI.

1.2 Element of human-defined objectives or parameters to be removed

We submit that, for similar reasons as given above, the element of human intervention may be too restrictive and that human-defined objectives/parameters may potentially be replaced by AI-defined parameters.

1.3 Element of 'without explicit programming' to be removed

We consider this element to be too restrictive and not technology neutral.

1.4 Element of 'influencing the environment they interact with' to be added

We consider that this element is an essential part of a definition of AI to ensure its relevance in a regulatory context. Without an element of influence or impact, the definition of AI remains abstract and theoretical.

Finally, we submit that any definition of AI, unless required by unique factors only present in Australia, should be as closely aligned with definitions adopted by international regulatory frameworks (e.g., the *EU AI Act*) to ensure compatibility and consistency for AI developers and deployers, which we consider to be key for ensuring competitiveness of Australia's technology sector in an international context.

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?

We consider that AI is regulated by several existing regulatory regimes, for example the *Australian Privacy Act 1988* (Cth), the *Competition and Consumer Act 2010* (Cth) and the *Online Safety Act 2021* (Cth), which are tailored to mitigate core risks associated with the use of AI, such as privacy, consumer and competition protections, and online safety. We are broadly supportive of the idea to rely on and build regulatory responses on existing regimes, as they are well established and technology neutral. However, we note that none of the aforementioned regimes address high-risk use cases of AI specifically and are currently not equipped to evaluate and, if required, ban AI applications with unacceptable risk levels. We consider that to the extent a separate regulatory regime were to be introduced to regulate the use of AI, such a regime should focus specifically on the testing, approving and banning of high-risk AI applications to close this gap.

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.

The Discussion Paper outlines existing non-regulatory initiatives, such as the DISR's AI Ethics Principles, pilot and Responsible AI Adopt Program, and the DISR via CSIRO's National AI Centre and Responsible AI Network. We acknowledge the ongoing development of these programs and initiatives and suggest, to the extent not already covered by these programs, the following:

(a) **government supported 'sandbox' environment** to develop and test new AI applications;

(b) **AI governance testing framework and toolkit** modelled on the "AI Verify" program adopted in Singapore to allow the testing of developed AI applications;[2] and

(c) **adoption of a government approved AI rating** (e.g., in relation to associated risk levels) and/or certification which signifies voluntary compliance with ethical standards, which may encourage businesses to participate in non-regulatory initiatives and build consumer trust.

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.

We consider that a designated AI board, composed of regulators involved in the regulation and administration of AI to date, such as the OAIC and ACCC, could assist in co-ordinating a unified across-government response in relation to the development, deployment and use of AI. Potential responsibilities could include the development of AI standards and best practices, issuance of guidance notes and hosting of AI related roundtables and consultations.

In particular, much like the guidance offered by the OAIC on privacy and consumer data right related matters, we propose that guidance in relation to any voluntary or regulatory AI framework would be an invaluable resource, in particular for SMEs, to implement responsible and safe AI practices.

We consider that the minimisation of patchwork solutions, and the establishment of a central AI body, particularly in the context of a regulatory AI framework, should be a key priority to further achieve across-government solutions and support responsible and safe AI practices in Australia.

[2] 'Make AI safe & beneficial for humanity', *AI Verify Foundation* (Web Page) <https://aiverifyfoundation.sg/>.

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?

Gadens considers the following governance measures and initiatives implemented by other jurisdictions may be adaptable and desirable for Australia:

Measure / initiative	Reasoning
European Union	
European AI Alliance – This is a public-private partnership initiative by the European Commission to establish a platform to facilitate collaboration, research and dialogue between stakeholders from various sectors to develop policies and strategies for AI.	Gadens considers that a partnership across a range of public and private sectors could be a beneficial initiative to allow wide stakeholder input and collaboration. It would offer an efficient platform for stakeholders to pool resources and jointly develop policies, strategies, as well as act as a forum for Australian regulatory bodies to test any proposed AI strategies or frameworks. A by-product of the Alliance is the High-Level Expert Group on AI, which is a subgroup of the Alliance which developed a detailed assessment list as a resource for businesses and organisations to self-assess the trustworthiness of their AI systems under development.
AI Watch – This is a knowledge service established by the European Commission to monitor AI development, uptake and impact of AI in the EU, as well as provide public access to resources on AI systems and development.	Gadens considers that a body such as AI Watch could assist the regulation and implementation of any AI policies and regulatory frameworks established in Australia. This may present itself in the form of an extended branch of the OAIC, or as a designated AI board, as proposed in section 4 of this submission.
Singapore	
AI Verify – This is a toolkit by the Infocomm Media Development Authority and the Personal Data Protection Commission which provides a testing framework for organisations to test their AI systems. Through standardised and technical tests and process checks, organisations can verify their AI systems and practices.	Gadens considers that a public toolkit or resource could assist businesses and organisations to better adopt AI systems in a measured manner by having a government-endorsed resource to self-assess the suitability of an AI system. This is similar to the detailed assessment list developed by the High-Level Expert Group on AI in the EU. Encouraging organisations to test their AI systems against a set of recognised principles will lead to greater transparency and uniformity. This may also protect consumers and provide a degree of public confidence that the AI systems adopted by Australian businesses and organisations must meet a

	certain minimum standard of security and quality.
Canada & New Zealand	
Risk-based and principles-based approach to classifying AI systems into risk categories Canada has issued a Directive on Automated Decision-Making which applies to federal government institutions and requires them to classify automated decision-making AI systems used to make administrative decisions into 1 of 4 risk categories based on the likely impact of the decision on individual rights, economic interests etc. Each risk category must satisfy a set of corresponding regulatory requirements. Similarly, New Zealand has also established an Algorithm Charter which sets out a risk matrix and classifies algorithms deployed by the New Zealand government into 3 risk levels. Various governmental bodies or government-affiliated institutions can become signatories to the charter.	Gadens considers that it would be beneficial to establish a risk matrix and set of principles (based on human rights, the Australian Privacy Principles and other core community interests) to guide the implementation of AI and automated decision-making systems by Australian Commonwealth government bodies and Commonwealth-affiliated institutions. This would ensure a level of consistency across Commonwealth institutions in relation to the use of automated decision-making systems and other AI systems. Further, it may assist to promote public confidence in the use of AI systems by the government. The Australian Government should also consider unifying a risk matrix and AI principles across all states and territories.

Target areas

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

Gadens considers that the public sector plays an important role in building public trust through the use of AI applications. As such, the public can expect that higher standards, whether voluntary or regulatory, should be met by the public sector.

We submit that while private organisations should strive to comply with the highest available standards, in particular in relation to ethical standards, mandatory compliance may negatively impact AI innovation and use in Australia. We consider that a mandatory and/or regulatory approach should be reserved for high-risk AI applications and to mitigate particular risk areas of AI, such as privacy, competition and protection of consumers and vulnerable people.

Gadens considers that industry specific codes of conduct and voluntary technical and ethical standards are the preferred tool to encourage AI best practices among private organisations while allowing them to remain competitive internationally.

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

Gadens submits that the following measures may assist the Australian Government in supporting responsible AI practices in its own agencies:

(a) **adaptation of a consolidated AI strategy for government agencies**;

(b) **implementation of a government-wide AI assurance and ethics framework** (similar to the NSW AI Assurance Framework and Ethics principles)[3]; and

(c) **a central or agency-specific AI register** which identifies AI applications implemented by relevant agencies, supplemented by risk assessments in relation to relevant use cases.

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.

We consider that generic solutions lend themselves to low risk and high-volume AI applications where generic solutions enable large audiences to benefit from AI applications.

While we acknowledge that technology-specific solutions require longer development and implementation timeframes, we consider that a combination of any one of the following factors may necessitate a departure from generic solutions:

(a) **complexity of an AI application**;

(b) **fast-evolving use case of an AI application**;

(c) **human oversight issues**;

(d) **high-risk impact of an AI application**; and/or

(e) **degree of regulation prevalent in the sector in which the AI application is predominantly deployed** (e.g., the medical sector).

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?**

We consider that transparency is equally important in both the development and the deployment phase of an AI application, and that all aspects of an AI application lifecycle are relevant to building public trust and confidence. In the development and testing phase we note that conformity with any mandatory regulation or voluntary conduct is a key factor. In relation to the deployment of AI applications we believe that consumers should be made aware of the fact that they are

[3] 'NSW Artificial Intelligence Assurance Framework', *Digital.NSW* (Web Page) <https://www.digital.nsw.gov.au/policy/artificial-intelligence/nsw-artificial-intelligence-assurance-framework>.

interacting with an AI application and, to the extent available and feasible in the context, be informed of the risk level associated with such use. Uniform risk levels/codes would be helpful in this context.

However, we note that to build public trust and confidence, it is not necessary to disclose all aspects of the development and testing process, full functionalities or all risks associated with an AI application. Transparency should balance the interests of businesses and consumers alike and should be used as a tool to both promote AI development and acceptance.

(b) **mandating transparency requirements across the private and public sectors, including how these requirements could be implemented.**

For the public sector we propose the following transparency requirements:

(i) **publication of applicable AI assurance and/or ethics frameworks and standards**;

(ii) **a central or agency-specific AI register** which identifies AI applications implemented by relevant agencies, supplemented by risk assessments in relation to relevant use cases;

(iii) **public consultations on implementation of large-scale, high-risk AI applications**;

(iv) **disclosure requirements in relation to the use of AI applications** (e.g., a uniform logo or symbol may be useful to flag an AI interaction to a user); and

(v) **disclosure requirements in relation to the risk associated with the AI application** (e.g., the risk report and/or assessment conducted by public organisations in relation to the AI application may be linked or a uniform risk code could be adopted).

For the private sector we propose the following transparency requirements:

(i) **publication of an AI policy** which outlines a private organisation's approach towards AI use (comparable to a Privacy or CDR policy);

(ii) **disclosure requirements in relation to the use of AI applications** (e.g., a uniform logo or symbol may be useful to flag an AI interaction to a user); and

(iii) **disclosure requirements in relation to the risk associated with the AI application** (e.g., high-level risk report for high-risk AI applications and disclosure of risk category (to the extent such is adopted) for low to medium risk applications).

10. Do you have suggestions for:

(a) **Whether any high-risk AI applications or technologies should be banned completely?**

Gadens considers that fast evolving AI use cases and open-ended designs of AI applications warrant the banning of very high-risk AI applications. Public trust will depend on ensuring that practices which are undesirable and detrimental to society are being banned.

(b) **Criteria or requirements to identify AI applications or technologies that should be banned, and in which contexts?**

We consider that a combination of any one of the following criteria or requirements, following a comprehensive risk assessment, may warrant the ban of an AI application or use case:

(i) **large scale and/or high complexity AI applications**;

(ii) **very high risk/impact expected** (eg, decisions that are threatening to life, welfare, financial well-being, freedom, particularly effecting vulnerable groups);

(iii) **irreversible/perpetual decisions**;

(iv) **applications with human oversight issues/high level of automation**; and

(v) **immoral** (social scoring for certain purposes) **or illegal** (facial recognition for tracking of the broader public without cause) **purposes**.

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

Gadens submits that the following initiatives and government actions can increase public trust in AI deployment:

(a) **government leading by example**, e.g., by adopting a mandatory government AI assurance and ethics framework (please refer to our answers above for detail);

(b) **government supported 'sandbox' environment** to develop and test new AI applications;

(c) **adoption of a government approved AI rating** (e.g., in relation to associated risk levels) **and/or certification** which signifies (voluntary) compliance with industry and/or ethical standards;

(d) **creation of an AI safety commissioner or portfolio** which allows users to complain about AI practices; and

(e) **consolidated AI education platform/centre** tailored to the general public (comparable to the DISR via CSIRO's Responsible AI Centre for industry).

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?

Gadens anticipates that an approach consistent with that taken by other jurisdictions, such as the European Union or Canada, with regard to banning of high-risk activities will have a positive impact on the trade with, and export of, AI applications by ensuring cross-jurisdictional acceptance and consistency. However, we note that any bans of high-risk activities should also be reviewed in light of decisions reached in other jurisdictions to

ensure Australia's tech sector is not unnecessarily disadvantaged in the international context.

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

Aligned with our submission to prioritise a gap analysis for existing regulatory regimes in relation to AI regulation, we submit that any changes made to the Australian conformity infrastructure should be informed by an infrastructure gap analysis prior to any changes being made. Further, Gadens considers it preferable to focus on changes which are necessary to protect against substantial AI risks (e.g., other than low level risks), as an overly prescriptive approach may not be technology-neutral and thus not future-proof, and stifle AI innovation.

Risk-based approaches

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

Gadens is generally in support of a risk-based approach to address the potential risks of AI. We note that many organisations will benefit from their expertise in relation to compliance with existing regulatory frameworks, e.g., the *Privacy Act 1988* (Cth) or the *Anti-Money Laundering and Counter-Terrorism Financing Act 2006* (Cth). We further consider that a risk-based approach will allow organisations to allocate resources effectively in respect of control implementation and integration of AI in their service delivery or process optimisation. We are not in support of prescriptive technology-focussed approaches as they will challenge SME AI developers and deployers and require careful monitoring and revision.

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

15.1 Benefits

We consider that a risk-based approach will:

(a) **enable companies to draw upon existing experience with risk-based assessments**;

(b) **allow for the timely adoption of low-risk AI applications**; and

(c) **prioritise resources** on the review and adoption of medium to high-risk AI applications.

15.2 Limitations

We consider that the main limitations of a risk-based approach are:

(a) **lack of historical data** to assist with accurate evaluation of risk events/issues (ie, likelihood and/or impact, which may result in low accuracy of risk assessments);

(b) **low accuracy of risk assessments** may impact the efficacy of controls; and

(c) **precautionary measures intended for medium to high-risk AI application may predominantly stifle cutting-edge and highly innovative uses of AI**.

15.3 Addressing limitations

Gadens considers that important strategies for overcoming any limitations associated with a risk-based approach are:

(a) **Increased focus on governing record-keeping and change management** – Record-keeping and change management activities associated with the implementation and use of AI applications should be a particular focus to overcome limitations and risks associated with incomplete datasets. Forums and roundtable discussions at industry or government level, as well as collaborations involving the sharing of governance mechanisms employed, may further facilitate risk-mitigation strategies and enable a timely adoption of a risk-based approach.

(b) **Testing of control effectiveness** – As companies harvest more data, control effectiveness should be tested and reviewed regularly, ensuring continual improvement. We consider the implementation of effective and frequent testing procedures a key driver of success for a risk-based approach.

(c) **Preserving flexibility in risk assessments** – While we consider improved governance and testing measures to be important risk assessment tools, we submit that the greatest benefits and innovations of AI may be stifled if risk-mitigation practices are inflexible and do not allow for progressive evaluations of AI applications. Supported by industry and government guidance, we consider that a degree of flexibility, which may not necessarily form part of traditional risk-assessment processes, is required to account for the fast pace in which AI evolves and is employed.

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?

No. While we consider that larger organisations may benefit from existing experience with risk-based assessments (e.g., in the context of AML/CTF compliance), risk-based assessments are scalable in that they are not solely based on the risk associated with an AI application but are also impacted by various other factors, such as an organisation's risk appetite. Further, potential disclosure obligations of AI developers/vendors and government guidance may assist smaller organisations to perform risk-based assessments, irrespective of size, AI maturity and resources.

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

17.1 Elements to be included

In addition to the elements listed in Attachment C, we consider the following control element appropriate to support a risk-based approach:

(a) **regular testing and review of control effectiveness** – To address the particular limitations of a risk-based approach (as outlined in 15.3(c) above), we consider that any controls implemented to address AI risks need to be tested and reviewed on a regular basis to ensure their effectiveness while enabling appropriate responses to innovative uses and/or developments in the area of AI.

17.2 Elements supported

Gadens supports all the elements listed in Attachment C. We further consider that impact assessments should be undertaken systematically at several points of AI integrations.

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?

As a legal services firm, Gadens places a particular emphasis on the protection of information – from both a privacy and information security standpoint. Thus, the adaptation of an AI-focussed risk-based approach would be heavily influenced by existing privacy risk and information security management processes.

Management processes are designed around information classification, which is impacted by:

(a) **legal, client or contractual requirements** that must be complied with;

(b) **value and criticality** to the organisation or their clients, and

(c) **sensitivity to unauthorised disclosure or modification**.

The type of information an AI application is required to process, store and transfer will be a determining factor for any AI risk-assessment to ensure the evaluation framework of the firm's information assets will align with the AI risk-based approach, streamlining the assessment process.

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

We consider that the control elements proposed in Attachment C will need to be evaluated in an application-specific context. While not all elements may lend itself to the evaluation of LLMs or MFMs, we consider that a risk-based approach, driven by frequently reviewed, organisation-internal risk frameworks, will enable risk-based assessments of emerging AI uses.

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?**

We consider that a risk-based approach for responsible AI should be mandatory for all developers and deployers (whether public or private) for high-risk AI applications. Generally, we would expect public organisations to lead by example, and a mandatory risk-based approach may be warranted irrespective of the level of risk involved, to facilitate public trust in AI applications.

For private organisations (developers or deployers) we consider that existing thresholds, such as annual turnover exceeding $100 million AUD, may be utilised to impose a mandatory risk-based approach irrespective of risk levels. Private deployers of AI below certain thresholds may benefit from a voluntary, industry-led framework. The regulation of small to medium size private developers of AI applications should be guided by further government and industry consultation,

taking into account the effect mandatory regulation would have on both innovation and trade.

Authors

Name	Title	Practice Group
Sinead Lynch	Partner	Intellectual Property and Technology
Freya Vom Bauer	Associate	Intellectual Property and Technology
Gemma Konsti	Legal Counsel	Office of General Counsel
Holly Chen	Risk and Governance Advisor	Office of General Counsel
Eliza Candappa	Risk and Compliance Officer	Office of General Counsel
Pearl Chen	Lawyer	Real Estate and Construction
Maggie Laing	Lawyer	Real Estate and Construction