Contact Simon Burns
T +61 2 9263 4776
sburns@gtlaw.com.au

Jen Bradley
T +61 2 9263 4288
jbradley@gtlaw.com.au

L 35, Tower Two, International Towers Sydney
200 Barangaroo Avenue,
Barangaroo NSW 2000 AUS
**T** +61 2 9263 4000 **F** +61 2 9263 4111
www.gtlaw.com.au

**4 August 2023**

Technology Strategy Branch
Department of Industry, Science and Resources

**Submission to the Department of Industry, Science and Resources: Safe and Responsible AI in Australia (Discussion Paper)**

## 1 Executive summary

Thank you for the opportunity to provide our views on the Safe and Responsible AI Discussion Paper published 1 June 2023 (the **Discussion Paper**).

We have set out in this document our responses to some of the questions posed by the Discussion Paper.

As an overview, our key comments are as follows:

- Particularly as we move toward decarbonisation, it is critically important for Australia to strike the right balance between managing risk and promoting our innovation and productivity agenda.

- Clear and efficient regulation, in parallel with sensible government initiatives (eg regulatory sandboxes and funding) is required to facilitate and encourage investment.

- In considering the best regulatory approach and policy setting it is important to acknowledge that:

    - Many of the harms or risks being experienced with AI today are the result of the lag between the incredible speed of AI evolution and the much slower development of organisational skills, understanding and capability surrounding AI. In other words, good intentioned organisations getting it wrong. An efficient and immediate regulatory response to this is enhanced guidance, support and toolkits to help organisations to "do the right thing"; and

    - Australia's existing multi-layer and technology neutral regulatory framework provides a solid base from which to consider management of potential harms across the whole AI lifecycle, but there are challenges and there are shortcomings within that existing set of regulations which need to be addressed to make sure they continue to be fit for purpose.

- Consequently, we agree with the approach contemplated by the Discussion Paper that there needs to be a detailed review of Australia's existing regulatory framework to identify where regulatory reform is required given current and expected AI capabilities.

- This review should be focused on:

    - identifying any gaps where existing laws do not function as intended in an AI context. For example, if there is a regulatory 'loophole" or similar where an organisation can avoid regulatory accountability by using AI in ways that are inconsistent with the regulatory intent;

    - identifying any gaps where existing laws do not provide coverage for an emergent harm or risk that arises as a result of new capabilities or shortcomings of AI that are inherently different to other technologies. For example, this may include consideration of issues surrounding mass surveillance and social manipulation, being use cases which are now possible due to evolution of AI capability. We note that this "gap" can often still be addressed through technology-neutral laws that are informed by the capabilities of emerging technology; and

    - identifying opportunities to remove regulatory burdens or controls that operate as obstacles to safe and responsible use of AI. The work of the National Transport Commission in relation to automated vehicles is a good example of what is required in this instance.

- The work outlined above requires a very detailed review undertaken by experts with deep domain expertise across each sector of our economy, and by each regulator within their remit. It is not a quick nor easy exercise, and it should not be rushed.

- The work also requires central regulatory and government co-ordination and capability to:

    - provide each sector and its regulators with access to AI expertise, common taxonomies, tools and frameworks;

    - ensure consistency with respect to regulatory objectives and approach, as well as output; and

    - ultimately help ensure that the final outcome and any resulting regulatory reform is harmonised and optimised wherever possible.

    In addition, this central function could address cross-sectoral issues, such as consideration of general purpose AI.

- We recommend that Australia continue to adopt a technology-neutral approach to regulation wherever possible. Creating new or overly prescriptive regulation or rules that apply to AI as a technology risks being unnecessarily rigid and inflexible and can create complexity and inconsistency when overlayed into our framework of existing laws. This approach can lead to very inefficient regulation that could quickly become out of date.

- Given the potential timeframe required to implement regulatory reform, we highlight that many of the risks posed by AI can be addressed through co-regulatory responses by assisting organisations to address gaps in their understanding of existing regulatory obligations in the context of AI, and increase their capability regarding AI. This can be a swifter initial response and risk mitigation approach.

Further detail regarding each of the above comments is included in our responses to specific questions posed by the Discussion Paper in the pages that follow.

## 2 About Gilbert + Tobin

Gilbert + Tobin is an independent Australian corporate law firm that is recognised as a leading transactions, regulatory and disputes law firm. We have one of Australia's most highly regarded corporate, commercial, technology and government practices, and are a legal adviser of choice for industry leaders and government on strategic, high profile and complex matters.

Our submission is guided by our experience in working with Australian government and business (many of whom operate in highly regulated environments such as financial services, healthcare and telecommunications) on the use of new technologies such as artificial intelligence and related matters, including personal information and data, intellectual property, cyber security and competition.

* * * * *

Thank you for the opportunity to provide our views on the Discussion Paper.

Please let us know if there is any further information or details that we can provide in respect of any of the above.

Yours sincerely
**Gilbert + Tobin**

**Simon Burns**
Partner
+61 2 9263 4776
sburns@gtlaw.com.au

**Jen Bradley**
Special Counsel
+61 2 9263 4228
jbradley@gtlaw.com.au

## Discussion questions

### *Question 1:*

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

1 Defining AI has been a common issue for regulators around the globe.[1] In considering whether to define AI, AI concepts or particular AI applications or techniques, it is important to consider why it is being sought to be defined, and in what context that definition is used.

2 We also note that regulating on the basis of a legal definition of AI and specific AI applications contrasts with the general approach to regulation in Australia which is technology neutral. As per our responses to other questions below, we support a sector-led regulatory and co-regulatory approach to regulating risks associated with AI that should, where possible, update and revise existing laws and remain outcomes-based and technology neutral. In this setting, it may not be necessary to define AI for statutory reform itself.

3 That being said, we acknowledge that a common understanding of AI and AI concepts across government would be useful guidance for a sector-led and technology neutral approach. Again, we reiterate that this should not be used to derive regulation targeted specifically at AI, except, potentially, in limited circumstances where a regulatory "gap" is identified that requires a technology specific, as opposed to a technology informed, response.

4 In forming a commonly understood definition of AI and AI concepts across government, the definitions should align to internationally used definitions and concepts to ensure global coherence, international competitiveness and ease of trade, but also to allow flexibility should industry understood definitions change over time.

5 It is also important to recognise that:

- AI is already, and will increasingly be, embedded in standard, off the shelf software products and solutions, including SaaS;
- in many instances the extent or nature of such embedded AI is not always fully transparent to acquirers or users of such products and solutions; and
- often the type of AI used is inherently no different to standard technology solutions.

6 The above will increasingly be the case if a broad definition of AI is adopted, as opposed to one that focuses more on the types of AI that give rise to novel risks that are different to other technologies – such as machine learning.

7 Consequently, whilst we appreciate the broad alignment to ISO definitions, we remain of the view that careful consideration is required as to where and how those definitions are used to ensure that unnecessary regulatory burden is not placed on organisations by making them apply risk assessments or similar to their use of AI technologies, which, depending on the technique or use case, may not be inherently riskier than any other technology solution or human process.

[1] For example, the European Parliament in its recent proposals to the draft EU AI Act proposed a narrower definition for AI that focused on a measure of autonomy (closer to the OECD definition) following criticism the original European Commission definition was too broad.

### *Question 2:*

### *What potential risks from AI are not covered by Australia's existing regulatory approaches?*

8 We recommend that the government undertake and co-ordinate a sectorial deep dive into Australia's existing regulatory framework as it applies to each sector - as the risks posed by AI, and the appropriate response to address those risks, are often highly context-specific.

9 Australia's multi-layer regulatory framework, which is primarily technology neutral, is already reasonably exhaustive in terms of seeking to address the types of harms that can occur as a result of use of AI, and regulators are increasingly turning to this existing regulatory toolkit to govern and regulate AI.

10 By way of example only:

- the Privacy Act regulates (amongst other aspects): the use of Personal Information as an input to AI solutions[2] including for training purposes; security in connection with such use;[3] the generation of new information about an individual via an AI solution (including regarding the accuracy of that information);[4] the use of biometric templates and other biometric information for biometric identification (which generally requires consent) and, if current reform proposals are implemented, requirements to conduct risk assessments and various controls and transparency requirements around substantially automated decisions that have a legal effect or similarly significant effect;[5]

- the Australian Consumer Law under the Competition and Consumer Act has:

  (i) prohibitions against misleading and deceptive conduct,[6] unconscionable conduct[7] and on making false and misleading representations,[8] which are relevant to how an organisation engages with consumers, collects data from consumers for an AI solution, the deployment of the AI solution, as well as any policy documents or statements that the organisation uses to explain or describe the AI solution and its outputs;

  (ii) strict liability offences on manufacturers to compensate consumers for personal injury and property damage that is caused by a 'safety defect'.[9] In an AI context, this *could* include a defect in the design, model, source data, or the manufacturing of the AI system, a failure to adequately test the system, including to address things such as bias or to make it sufficiently secure against cybersecurity attacks; and

  (iii) statutory guarantees in relation to goods and services.[10] For example, the statutory guarantees provide that any goods supplied to a consumer must be of acceptable quality, fit for purpose and match their description, and, in respect of services, that the services are rendered with due care and skill and are fit for any stated purpose. Poor quality outputs or outcomes (e.g. inaccurate results) by an AI solution *may* result in it breaching these guarantees depending on the surrounding circumstances and the product delivery model (eg as a good or service) leaving a

[2] See for example, Office of the Australian Information Commissioner, *Australian Privacy Principles* (as at 31 July 2023) Principles 1, 3, 6 and 11

[3] Ibid, Australian Privacy Principle 11.

[4] Ibid, Australian Privacy Principle 10.

[5] See for example, Attorney-General's Department, *Privacy Act Review* (Report, 2022) Proposal 19.

[6] *Competition and Consumer Act 2010* (Cth) Sch 2 ('*Australian Consumer* Law') s 18.

[7] Ibid 20-22A.

[8] Ibid 29-37.

[9] Ibid Part 3-5.

[10] Ibid Part 3-2.

supplier liable to provide certain remedies to consumers, such as a refund, replacement or reperformance; and

- the various State and Territory anti-discrimination laws have controls to prohibit discrimination against any person on the basis of protected attributes. This is relevant both to the use of the AI solution (e.g. because it is not accessible to a person with a protected attribute, such as a disability) or because the outputs or outcomes of the system are discriminatory. Anti-discrimination laws therefore regulate the data collection and ingestion stage, as well as training, testing and eventually ongoing monitoring of outputs by requiring organisations to assess whether there are, in fact, any biased or discriminatory outcomes in respect of protected attributes.

11 However, whilst the regulatory intent of these existing laws with respect to protection of privacy, safety, consumer protection and anti-discrimination align with many of the potential harms that arise in relation to AI, there are potential shortcomings and gaps that require examination to ensure that AI does not, through a loophole, fall outside these existing regulatory regimes and that the regulatory remedy or response is adequate in an AI context. In other words, to ensure that our existing laws remain fit for purpose for the technological advancement of AI,.

12 For example, there is a potential gap in existing laws in relation to our product liability laws, and what is known as the 'state of the art defence' or 'development risk defence'. Under the Australian Consumer Law (Part 3-5), a manufacturer of an AI system is strictly liable to compensate consumers for personal injury and property damage that is caused by a 'safety defect' in the AI system, being a defect in goods that are not as safe as a person is entitled to expect that they are. However, under the development risk defence, a manufacturer could disclaim liability if they establish that the state of scientific or technical knowledge at the time when the good was supplied was such that the manufacturer was unable to discover the defect. In other jurisdictions, such as the EU, the defence has been called into question in the context of AI,[11] as it could be used by a manufacturer of an AI system where the AI outputs that caused the harm were unpredicted, or as a result of the AI system self-learning. The ACCC could be charged at looking at whether this is a potential gap, or whether regulatory guidance could be issued to clarify what would be considered a 'safety defect' in the context of self-learning systems – for example, it may be considered a 'safety defect' and a manufacturer not able to rely on the 'state of the art defence' if there was an absence of review, auditing or other in-built oversight mechanisms in the AI system to mitigate against unpredicted or self-learning behaviours that may cause damage in the future.

13 Further, the product liability regime will only apply to AI solutions which are considered 'goods' under the ACL (which is broadly defined to include computer software).[12] However, the supply of software or digital systems such as AI would be considered a service in certain contexts, depending on the nature of the software or system and the distribution model in question. Given the possible ways in which AI may be used, consideration should be given as to whether the there is a gap under the product liability regime if certain use cases of AI are considered a service.

14 These types of "gaps" are arising in two main ways: Firstly, AI and its capabilities and shortcomings are (in some, but not all cases) inherently different to other technologies, and bring to the fore the potential for new harms and risks as well as the opportunity for new use cases (both good and bad). For example:

- The dynamic and self-learning capabilities of certain AI systems means that AI systems may act in unintended and unprogrammed ways; creating difficulties with error identification and operational risk management, and making it more difficult to assign legal liability and accountability in some contexts within our legal system.

[11] See for example, *New liability rules on product and AI are encouraging but need improvement* (Blog Post, 4 October 2022). <https://www.iccl.ie/news/new-liability-rules-on-product-and-ai-are-encouraging-but-need-improvement/>.

[12] Ibid (n-6) 2.

- The level of sophistication and capability of some AI systems may mean that use cases which were previously impossible or impractical to deploy now a possibility, such as mass surveillance and social scoring. This could give rise to a regulatory "gap" that needs to be addressed potentially via reform of existing surveillance or privacy laws.

- Discrimination laws that regulate AI are only for certain protected attributes. However the risk of AI systems entrenching and amplifying bias within our society based on use of biased training data is a new issue that requires new thinking, including because this bias can be based on attributes not currently regulated. We are not suggesting we need to completely reform discrimination laws specifically for AI nor "reinvent the wheel" on anti-discrimination legislation, but we agree with the Discussion Paper that we need to consider this potential societal harm carefully.

15 Secondly, the exponential speed of AI development and the step change that we are in the midst of means that organisations often lack the capacity, capability, guidance and tools to assist them to safely and confidently deploy AI or manage their existing regulatory obligations in the AI context. This is a slightly different issue, and we consider that the primary regulatory response for this should be focused on government assistance and guidance.

16 For example, the often poorly understood ability of generative AI to hallucinate is obviously a risk that has tripped up many an organisation already. This is a clear example of the failing of existing risk and quality assurance frameworks and lack of organisational capacity and capability to understand and govern AI use. Given this typically gives rise to risks and negative consequences for the user of that AI solution (as opposed to solely negative externalities), an efficient and effective regulatory response can be through issuing guidance and developing toolkits and frameworks for organisations to understand and manage this risk. However, consideration should also be given to the particular use cases where the extent of the negative externalities mean that regulatory intervention may be appropriate.

17 We also note that, whilst the Discussion Paper does not seek to address intellectual property issues, there is a need to examine whether or not our copyright system remains fit for purpose in the context of generative AI, and we note the separate initiatives in this regard as referred to in Attachment A of the Discussion Paper.

***Do you have suggestions for possible regulatory action to mitigate these risks?***

18 Given the potential timeframe required to implement regulatory reform, we highlight that many of the risks posed by AI can be addressed through co-regulatory responses by assisting organisations to address gaps in their understanding of existing regulatory obligations in the context of AI, and increase their capability regarding AI. This can be a swifter initial response and risk mitigation approach. Please refer to our answer to Question 3 for more details.

19 Notwithstanding this, ultimately addressing the types of harms and risks referred to above requires regulatory reform based on thoughtful and nuanced analysis and sectorial and industry expertise, involving multiple disciplines, including policy, tech, legal and civil society.

20 Ideally, this regulatory reform would leverage our existing set of technology neutral laws, and make sure they are fit for purpose having regard to AI. This includes a need to:

- consider what laws may be required to prevent or mitigate ill-intentioned actors from being able to leverage the new capability of emergent AI models for inappropriate or intentionally harmful purposes;

- ensure that where harmful outcomes do occur, the law is able to adequately respond to ensure there is appropriate accountability across the AI supply chain and that redress or rectification is available;

- cross check our existing set of technology neutral laws to ensure they can be applied effectively and that there are no loopholes, gaps or inconsistencies that mean AI implementations escape the regulatory intent;

- ensure regulators have the necessary mandate, capability and resources to enforce the law and rules in their sector in the context of AI; and

- identify opportunities to remove regulatory burden or controls that operate as obstacles to safe and responsible use of AI.

21 This is not an easy task and requires a contextual understanding and analysis of AI use cases in each industry.

22 We caution that a "one size fits all" economy wide AI-specific law risks duplication with our existing set of laws, inconsistency across regulation and unintended negative consequences by either over or under regulating AI use in certain sectors or contexts. Instead, we consider that review and analysis should be conducted on a sector-by-sector basis, where specific regulators are charged with looking at context-specific responses to particular use cases, or in responding to specific gaps in existing laws within their sector.

23 As highlighted in the Discussion Paper, the range of contexts in which AI can be used and for different purposes may often necessitate context-specific responses to regulating AI. A key advantage of a sector-led approach is that both:

- the risks associated with AI systems and their resulting potential impact; and

- the appropriate steps to mitigate those risks,

24 is often dependent on the AI technique used and/or the environmental context in which it is being deployed, where context means not only the purpose for which an AI system will be used (including balancing its potential benefits), but also the legal context in which it will operate. For example, the type and level of risk and harms associated with use of a machine vision system used in cancer diagnosis are different to a machine vision system used in logistics. Within an individual sector such as financial services, the risks and appropriate mitigants for biometric identification systems used for KYC checks warrant a different response to those in fraud prevention or detection. Further, different considerations should be applied to questions of maintaining human autonomy rather than outsourcing human tasks to machines in different areas, for example back-end process automation vs creative/copyright industries.

25 Horizontal, sector wide rules risk missing these nuances and may lead to unintended consequences rather than addressing the real concern. A sector-based approach allows flexibility for existing regulators to address and manage the impact of the use of AI in their sector, including in response to new technological developments and specific AI use-cases, applying their contextual expertise to consider any gaps in existing sector frameworks, consider questions of the appropriateness for AI to perform activities traditionally performed by humans in their sector, and impose the appropriate level of scrutiny.

26 Further, a sector-led response has the advantage of ensuring that the same regulators who have experience in managing and enforcing manual or human-based processes in their sectors, also regulate the outcomes generated by AI. Thus avoiding anomalies that may arise if a certain act is only an issue or treated differently if it uses AI to generate an outcome, as opposed to relying on a manual or human based process. Such a two-prong approach may give rise to other unintended consequences such as reduced investment and adoption of AI and its potential benefits, if the human-based process is not so regulated.

27 Accordingly, where any gaps are identified in a sector-led review, any law reform response should, as far as possible, be technology neutral.

28 We have already seen the beginnings of a sector-led gap analysis in Australia, for example by the National Transport Commission (NTC) in relation to its work on autonomous vehicles, where the NTC found more than 700 barriers to the deployment of autonomous vehicles in Australian laws, which are designed around vehicles having a human driver.[13] There is also other work underway, for example proposals put forward in the Privacy Act Review Report to review the *Privacy Act 1988* (Cth) in relation to practices that may touch on the use of AI and a potential reconsideration of exceptions under the Act,[14] together with reforms in other areas such as those highlighted in the Discussion Paper.

29 Regulators themselves also need the budget, the capacity and capability to undertake this exercise and there would be benefit in having a centralised AI capability that individual regulators could access.

30 Further, we recognise that a sector led approach could lead to less oversight of general-purpose AI that impacts multiple sectors, or lead to an inconsistency in approach which can create regulatory complexity and inefficiency, particularly when AI solutions can be used across sectors. In addressing this risk, we see the benefit of the introduction of centralised functions and mechanisms (see Question 3 for further commentary).

***Question 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.***

31 It is important to recognise that the majority of organisations involved in AI development and use are well intentioned and want to do the right thing, but lack practical capacity, capability and tailored guidance to ensure AI is being implemented safely and responsibly in line with AI Ethics Principles and within the bounds of their existing legal frameworks.

32 While a regulatory gap analysis does need to take place (as outlined above in Question 2) this will take time. In our view, an initial regulatory response should be focused on preventing or mitigating inadvertent and unintentional, but harmful outcomes by assisting organisations to:

- **develop capacity and capability regarding AI**: through funding, training and skills development;
- **understand specific risks in their field and to comply with existing regulatory frameworks in the context of AI:** through issuing tailored regulatory guidance; and
- **implement processes that guide and lead to responsible AI use**: such as through leveraging governance, risk management and assurance frameworks and technical standards.

33 For this, skills development at both the organisational and regulator level will be essential. We also see the importance of other co-regulatory tools, such as the use of sandboxes, such as ASIC's enhanced regulatory sandbox, which can facilitate the trialling of new innovations in a controlled environment and under regulatory supervision and feedback. Regulatory sandboxes can also help inform regulatory gap analysis (see Question 2) by highlighting how existing requirements might need to be clarified or modified, or where necessary, new laws be created.

---

[13] National Transport Commission, *The regulatory framework for automated vehicles in Australia* (Policy Paper, February 2022).

[14] Such proposals under the *Privacy Act Review Report* include increased transparency requirements and individuals rights related to 'targeting', transparency requirements and individual rights associated with automated decisions which have a legal or similarly significant effect, together with more general proposals such as the requirements that use of personal information be 'fair and reasonable', the requirement for privacy impact assessments for high risk activities, the expanded definition of personal information, and expanded individual rights.

34 Other co-regulatory responses include the leveraging of standards, assurance frameworks and certification schemes under the guidance of regulators on their application, to assist businesses in the governance and mitigation of AI risks and the promotion of safe and responsible AI in design, development, testing, implementation and maintenance.

35 As referenced in the Discussion Paper, global standards organisations, such as the National Institute of Standards and Technology and the ISO/IEC (JTC 1/SC 42) have developed and are further developing standards and frameworks to assist organisations with AI, including AI risk management standards (see for example the NIST AI Risk Management Framework[15] and ISO/IEC 23894:2023[16]) and AI data quality standards are currently under development (ISO / IEC 5259-4). CEN and CENELEC is also developing standards for the proposed EU AI Act (CEN-CLC/JTC 21).[17] At a local level, the NSW Government has also introduced the NSW Government AI Assurance Framework[18] to assist NSW government agencies (and suppliers to NSW Government agencies) to design, build and use AI-enabled products and solutions appropriately, safely and in line with ethical principles.

36 In this regard, risks associated with AI could be managed in a similar manner to the way in which privacy or cybersecurity is regulated, where industry-based approaches such as technical standards or third party certification and audits are used to support and compliment outcomes-based regulation. For example, regulation sets the requirement that an organisation must 'take reasonable steps'[19] or implement 'commensurate'[20] security controls against loss, misuse, disclosure etc of information. The regulations themselves do not set out what 'reasonable' or 'commensurate' steps are, but adopting an industry standard of technical security practices is then one way that organisations seek to meet their outcome-based compliance obligation.

37 There are a number of advantages to this approach, including that:

- legislation remains largely technology neutral and outcomes based, avoiding the anomalies raised above, while promoting risk mitigation measures in the AI development process;
- it allows for greater flexibility in response - what is appropriate assurance for a particular use case or a particular organisation (be it large or small) is able to flex and adapt more readily than if it was enshrined in legislation; and
- regulatory guidance and standards are more readily able to be reviewed and updated to (i) address new risks and new technologies than the traditional lag time in legal and regulatory development, and (ii) to reflect any judicial interpretation as to how the law applies to AI.

38 Further where international standards and assurance models are leveraged, this will allow Australia to align to internationally accepted approaches, promoting global interoperability (and consequentially ease of trade).

39 Leveraging these co-regulatory responses will in turn be a driver for encouraging ongoing innovation and investment in AI, and ensure better outcomes for public consumers.

---

[15] National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, available: https://www.nist.gov/itl/ai-risk-management-framework

[16] International Organization for Standardization, *ISO / IEC 23894:2023 Information technology – Artificial Intelligence – Guidance of Risk Management*, available https://www.iso.org/standard/77304.html

[17] See, *CEN-CENELEC JTC 21 'Artificial Intelligence'* (Webpage) <https://www.cencenelec.eu/areas-of-work/cen-cenelec-topics/artificial-intelligence/>.

[18] NSW Government, *NSW Artificial Intelligence Assurance Framework* (Report, March 2022).

[19] Office of the Australian Information Commissioner, *Australian Privacy Principles* (as at 31 July 2023) Principle 11.

[20] Australian Prudential Regulation Authority, *Prudential Practice Guide: CPS 234 Information Security* (June 2019) Para 21.

***Question 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.***

40 As mentioned in our response to the questions above, in order to address the risks in a sectorial approach, we see the benefit of including centralised mechanisms such as a central federal body, committee or multi-regulator working group(s) in the overall regulatory response.

41 In this regard we see aspects of the proposed UK approach as a reference model. Under the UK approach,[21] existing regulators are charged with reviewing laws and giving organisations within their respective field guidance to assist them to implement a common cross-economy set of responsible and ethical AI principles, while creating a single central committee or body that is charged with:

- **coordination and harmonisation:** across regulators and state and federal governments to ensure consistency, address regulator overlap, minimise duplication of effort and promote alignment in applying regulatory frameworks and guidance across industries in the context of AI. Further, coordination may also involve a consideration of balancing competing interests and trade-offs of different regulators in the AI space, for example fair competition against data protection;
- **support regulators technical skills and industry dialogue:** by providing technical skills training and advice on AI to regulators. This could leverage and be supported by the advice of expert advisory groups; and
- **monitor cross sector AI:** to identify and review cross-cutting AI risks, including in response to general-purpose AI and responses which could lead to the application of specific legislation if gaps are identified.

42 We note that other regulatory coordination bodies have been set up for other cross-industry issues, such as the Digital Platform Regulators Forum between the ACCC, ACMA, OAIC, and the eSafety Commissioner.

***Question 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?***

43 As set out in Question 4, we see the UK's approach of leveraging existing regulators to regulate (including through AI related regulatory reform) the use of AI in their sector against a common set of principles, supported by a central coordination and oversight function, as a relevant, adaptable and desirable reference model for regulatory reform in Australia. See also the advantages of a sector-led approach as set out in Question 2. We also consider that this type of pro-innovation approach is consistent with Australia's innovation and productivity agenda.

44 In contrast, while there are elements of the EU approach that are sensible depending on context, and are common to other assurance and risk management approaches, there is a risk that the EU approach has a high regulatory burden and is less flexible and efficient.

45 While, at a macro-level, the EU AI Act[22] proposes to regulate AI systems proportionately based on risk, it is reasonably prescriptive and, if adopted in Australia, risks overlap with other existing laws (for example, the Privacy Act) and unintended negative consequences. We also note the

[21] Department of Science, Innovation & Technology, *A pro-innovation approach to AI regulation: CP 815* (Report, March 2023).

[22] European Commission, *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence*, COM/2021/20 (EU AI Act).

obligations and assurance processes imposed on 'high risk' systems are broad and are, to a large extent, enshrined in legislation. The ability to amend the lists of use cases within each risk category will either require amendment to the EU AI Act (in the case of 'unacceptable risk') or require a delegated act, subject to certain conditions (in the case of 'high risk'). Amending the rules that apply to each risk category will also require legislative amendment (as highlighted by the introduction of specific rules by the European Parliament for generative AI that did not fit neatly within the existing tiered classification system). It is therefore not easily able to flex with changes in values or technological development.

46 As has been evident from the slow pace of negotiation of the draft EU AI Act, there are difficulties in defining AI, assigning and future-proofing general risk categories and therefore what and in what circumstances AI systems should be regulated or banned. Some of these challenges result from the economy-wide approach adopted.

47 Ultimately however, we agree with the comments in the Discussion Paper that whatever approach is adopted in Australia will have to factor in and have some alignment to international approaches given the global nature of AI, and the extra-territorial application of other proposed approaches, such as is the case with the EU.

***Question 6 and 7:***

***Should different approaches apply to public and private sector use of AI technologies? If so, how should the approaches differ? How can the Australian Government further support responsible AI practices in its own agencies?***

48 As with the private sector, the risks and benefits associated with the use of AI by the public sector will be context specific, depending on the government activity involved. While AI holds substantial promise for how to improve the delivery of government services, when considering both the safe and responsible use of AI by government, there are likely to be a higher number of high-risk use cases for reasons such as:

- the public do not have a choice when it comes to interacting with government services;
- government decision-making (whether made by AI or a human decision maker) will often impact a person's legal rights or similarly significant rights;
- the sensitivity of government data as an input to AI;
- government activities (including digital transformation projects) are funded by taxpayer dollars; and
- community expectations as to government behaviour, and the importance of transparency, fairness and accountability in government activity.

49 If a sector-led approach is followed, then the approach to public sector use of AI can be considered as its own sector within which the gaps, risks and appropriate mitigants, and guidance when it comes to AI can be appropriately tailored to reflect the unique characteristics of the public sector.

50 Further we support central coordination when it comes to government policy and guidance relating to AI use in government, beyond any agency specific guidance. We note at a Commonwealth level, the Digital Transformation Agency and the Department of Industry, Science and Resources are collaborating and investigating the need for whole-of-government policies and standards relating to use of AI in government. We support such collaboration extending to State and Territory government also.

***Question 8:***

***In what circumstances are generic solutions to the risks of AI most valuable? And in what circumstances are technology-specific solutions better?***

51 As per our responses in other sections, any regulatory response to AI should ideally be technology neutral and outcomes based to allow them to be flexible and readily adaptable.

52 Too much focus on the technology gives rise to a risk that the laws need to be constantly updated to address new AI techniques and applications as they are being developed. Given the pace of this development, this will naturally result in significant regulatory lag.

53 There is also risk that creating an AI-specific regulation (whether generally or for a particular AI application) may also:

- create duplication and overlap between existing general and sector-specific regulation which can create a complicated compliance burden for organisations, which may ultimately lead to less compliance, and ineffective regulation, whilst also stifling innovation; and

- lead to anomalies if a certain act is only an issue, or treated differently, if it uses AI to generate an outcome, as opposed to relying on a manual or human based process. Such a two-prong approach may give rise to other unintended consequences such as reduced investment and adoption of AI and its potential benefits, if the human-based process is not so regulated.

54 Technology-specific regulatory requirements should be limited to those scenarios where a gap is identified in existing law that is unable to be addressed following a technology-neutral approach.

***Question 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.***

*When transparency is most critical*

55 We do not agree with the implication that transparency (awareness) is always required across the entire AI lifecycle. Rather, the proportional importance of transparency (i) for a particular AI system, (ii) in what parts of the lifecycle of that AI system, and (iii) to whom, will depend on the type of AI being used, the use case of that system and the context in which it is used.

56 This is particularly the case if a broad definition of AI is adopted which captures a range of AI technologies that do not give rise to inherently different risks than other types of technology.

*Mandating transparency*

57 Similar to our other comments on a sector-led approach, in our view it is a decision for individual regulators to consider when and to what extent transparency principles are important and how they should be implemented (whether by guidance or mandated through regulation) to address potential risks and public trust within their sector.

58 There are already examples of a movement towards mandating transparency in certain circumstances. For example, amendments proposed as part of the Privacy Act Review Report include increased mandated transparency requirements for private entities and commonwealth agencies governed by the *Privacy Act 1988* (Cth). Privacy collection notices will be required to include information about the circumstances of collection, use and disclosure of high risk privacy activities.[23] Many AI systems that rely on personal information are likely to fall within the definition of a high-risk privacy activity[24] and consequently subject to the increased transparency requirements in privacy collection notices. Further specific transparency requirements have been included for collection, use or disclosure of personal information for the purposes of substantially automated decision making with legal or similarly significant effects, and similar transparency requirements are proposed for AI systems that use information about an individual (including de-identified and non-identifiable information) to track, profile or deliver targeted content to individuals.[25] The proposals also lead to meeting explainability principles, with new rights for individuals to an explanation or summary of what an organisation has done with their personal information in any circumstance,[26] together with a right to obtain meaningful information about how an automated decision with legal or similarly significant effects is made.[27]

59 These proposals in the Privacy Act Review Report illustrate a risk-based approach to applying principles of transparency and explainability within the OAIC's remit.

60 Whether transparency is required, at what stage and to whom will depend on the contextual circumstances. While transparency in certain cases may be important (such as deep fakes), in other low impact settings it may have negative effects, such as creating a poor user experience or high compliance burden, leading to decreased deployment and uptake, or decreasing the effectiveness of the tool (particularly if it is being used to mitigate criminal or harmful behaviour, such as fraud detection). Further, regulators may be of the view that regulatory guidance on transparency requirements, including how transparency with respect to AI fits within other existing mandated requirements, is sufficient.

***Question 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?***

61 We caution against blanket bans in respect of particular AI technology or applications. This is because the risks or potential harms posed by a particular AI technology or application are contextual. It is the use case that should be examined rather than the AI technique / application itself. Further, hard-coding prohibitions on particular use-cases in legislation (particularly broad or general descriptions) is highly inflexible, particularly as societal attitudes change, technology improves and develops or where appropriate guardrails, mitigations and human oversight can be implemented. A blanket ban risks unforeseen negative impacts, such as prohibiting possible beneficial uses of that AI technology or application in the future or in particular benign settings.

62 We also note the role that existing laws in Australia, in particular privacy laws, anti-discrimination laws, consumer laws and criminal laws, play (or can play) in regulating undesirable or high risk activities.

---

[23] Attorney-General's Department, *Privacy Act Review* (Report, 2022) Proposal 10.2.

[24] Ibid pg. 124, which sets out a list of indicative list of 'high-risk' activities which included the collection, use or disclosure of sensitive information or children's information on a large scale, the use of biometric templates or biometric information.

[25] Ibid Proposal 20.9.

[26] Ibid Proposal 18.1.

[27] Ibid Proposal 19.3.

63 Under the Privacy Act Review Report, it is further proposed that the *use* (not just collection) of personal information must also be fair and reasonable[28] and that privacy impact assessments will be required for all high privacy risk activities[29] - the outcome of which may be the identification of risks such that the particular activity should not proceed.

64 Requirements such as these in existing laws will generally require organisations and agencies to consider the fairness, reasonableness and risk associated with that particular AI use-case. The application of these considerations is able to adapt to then-current social and ethical standards, be informed by specific regulatory or agency guidance, as well as the technical capacity, accuracy and applications of AI as they change over time.

65 Of course, certain use cases warrant more consideration for prohibition, particularly in the context of new AI capabilities that may enable large scale societal harms, but these use cases need to be considered within existing frameworks to consider whether there is in fact a regulatory gap that needs to be addressed.

***Questions 14 – 17:***

***Do you support a risk-based approach for addressing potential AI risks? If not, is there a better approach? What do you see as the main benefits or limitations of a risk-based approach? How can any limitations be overcome? Is a risk-based approach better suited to some sectors, AI applications or organisations than others based on organisation size, AI maturity and resources? What elements should be in a risk-based approach for addressing potential AI risks? Do you support the elements presented in Attachment C?***

66 We are generally supportive of a risk-based approach, which focuses on high-risk uses of AI and not low-risk uses. This has the advantage of focusing on the areas where there are most risks of harmful outcomes while not imposing an undue compliance burden on organisations which may ultimately lead to less compliance, ineffective regulation, and a stifling of innovation.

67 However, as per our earlier comments, it is important that risk categories and the appropriate response requirements for each category are not overly rigid, and the risk threshold is set at the right level. As has been evident from the slow pace of negotiation of the draft EU AI Act, there are difficulties in assigning and future-proofing general risk categories and therefore what, in what circumstances, and how AI systems should be regulated.

68 AI is often an iterative technology, constantly learning in response to outputs and new techniques are constantly being developed. Further, the risks associated with AI systems, and the appropriate steps to mitigate those risks, are dependent on the technological and environment context in which they are being deployed and their outcomes. This is relevant for both providers and deployers of AI systems. Accordingly, while we are generally supportive of the principles contemplated by Attachment C, when and how they should be applied will depend on the context of the system and we caution against overly prescribing this in legislation. Where possible and practicable, Australia should leverage international standards and assurance in order to align to internationally accepted approaches, promoting global interoperability (and consequentially ease of trade and reducing compliance burden for global businesses).

69 This leads to support for a sector led, principles and outcomes-based approach that focuses on high risk uses but allows the appropriate risk settings and response to adapt and change based context. This is inherently difficult with a horizontal, sector wide, rule-based approach which risks being unnecessarily rigid, and may stifle innovation by forcing AI applications and assurance processes into static regulatory categories which are less able to adapt.

[28] Ibid Proposal 12.

[29] Ibid Proposal 13.1.