LEGALVISION®

26 July 2023

Technology Strategy Branch
Department of Industry, Science and Resources

By email: DigitalEconomy@industry.gov.au

Submission: Discussion Paper on 'Safe and Responsible AI in Australia' opened on 1 June 2023

Thank you for the opportunity to respond to the Discussion Paper on 'Safe and Responsible AI in Australia', opened on 1 June 2023, from the Department of Industry, Science and Resources (**Discussion Paper**).

LegalVision is a tech-driven commercial law firm that operates primarily online. The majority of our clients are start-ups and small to medium businesses, many of which are companies operating in the artificial intelligence (**AI**) space that are keen to see a considered and balanced regulatory regime deployed in Australia.

In responding to the Discussion Paper, we welcome any changes proposed by government that support the following objectives:

- Proportionately regulates AI development, deployment and use without stifling innovation, primarily using a risk-based approach
- Is flexibly defined so as to capture the many possible use cases of AI, both in the present and in the future
- Provide appropriate stakeholder protection
- Plugs the gaps in existing regulatory frameworks concerning AI

Our responses to the specific questions in the Discussion Paper follow in Annexure A.

Please direct any questions you may have to:

Joel George | joel.george@legalvision.com.au

or

Robert Chen | robert.chen@legalvision.com.au

Yours sincerely,

LegalVision ILP Pty Ltd

LegalVision

Annexure A – Responses to Specific Questions

OVERVIEW: Proposed Regulatory Action

Answer

The regulatory framework of artificial intelligence (**AI**) in Australia should take both a regulatory and a non-regulatory response. We have set out each legal "topic" relevant to AI against their proposed response in the tables below.

Topic	Proposed regulatory response
AI generally	AI specific legislation (**AI Legislation**) should be passed that adopts a risk-based approach. A high level proposal has been set out in **Annexure B**. **Risk-based approach** A risk-based regulatory approach for AI in Australia would involve assessing the risks associated with different AI systems and then tailoring the level of regulation to those risks. This approach would allow Australia to regulate AI in a way that is flexible and proportionate, without stifling innovation. There are a number of benefits to adopting a risk-based regulatory approach for AI. First, it would allow Australia to focus its regulatory resources on the AI systems that pose the greatest risks. This would help to ensure that Australia is addressing the most serious risks associated with AI, while also avoiding overregulation of AI systems that pose low to medium risks. Second, a risk-based regulatory approach would allow Australia to be flexible and adaptable in its regulation of AI. By focusing on the risks posed by an AI system (i.e. a somewhat functional assessment of the AI system), the AI Legislation would be well equipped to account for the various use cases of AI systems, both now and in the future. Third, a risk-based regulatory approach would be consistent with other countries who share similar values to Australia. Under the Western, Educated, Industrialised, Rich, and Democratic ("WEIRD") framework, a typology of countries based on how Western, Educated, Industrialised, Rich, and Democratic they are, Australia, the EU and Canada all fall under the same category. South Korea is typed similarly to Australia too, though only less "Western". The AI Legislation will operate similarly to the AI regulatory framework in the EU, Canada and

	South Korea. These countries present good case studies as they still show relatively strong AI development whilst requiring AI systems that pose high risks to comply with a number of requirements, such as transparency, explainability, and robustness. By adopting a risk-based regulatory approach for AI, Australia would be aligning its regulatory framework with standards and best practices of other countries with similar values. This would help to ensure that Australia's regulatory framework is seen as credible and legitimate by businesses and other stakeholders. This will also help promote ease of business for AI business which, given the nature of AI systems, are likely to operate internationally. **Regulatory force** In our view, the AI Legislation should have the force of law through legislation and government enforcement; it should not be left as a voluntary or self-regulatory tool. Reasons for this include: 1. Voluntary or self-regulatory tools can be less effective than government regulation because they rely on the initiative and cooperation of businesses and organisations. 2. Voluntary or self-regulation by businesses and organisations will come at costs to businesses and organisations which decreases the likelihood of sufficient self-regulation, particularly, in the current economic climate. 3. AI can have significant potential positive or negative effects on all aspects of society, depending on its use. Without firm regulation, there is a possibility that AI can be used for significantly negative uses and result in significant harm to certain aspects of society. 4. Government regulation in AI will grant greater clarity and certainty to businesses in the AI space. 5. Government regulation will grant the public greater protections and potential remedies against harms caused by AI use and thereby increase public confidence. 6. At this stage, we do not consider the imposition of a licensing regime to be appropriate as it would inhibit or slow down any AI development.
Regulatory Sandbox	A regulatory sandbox for AI should be put into place in Australia for businesses and organisations to innovate. This sandbox should be administered by a regulatory body (which may be an existing regulator or a newly created body specific to

	AI Legislation). Many of our clients have found regulatory sandboxes, like the financial services sandbox administered by ASIC, to be particularly useful. This is especially the case because: 1. many of our clients are startups who are still developing and designing a product/service; and 2. the financial services regulatory framework adopts a functional approach (which is not dissimilar to the AI Legislation proposed). We note that jurisdictions like the UK and the EU are also contemplating putting AI sandboxes in place. These sandboxes will be useful as it will provide an environment free from regulatory risk, for businesses to develop and test AI systems. It will help encourage development and innovation of AI technology. It would also provide invaluable information which can serve as feedback to any AI regulation by placing the regulator administering the sandbox at the coalface of AI technology development. An exception should be made for the AI uses under the “Unacceptable” risk profile in the AI Legislation which, due to their nature and potential impact, should remain prohibited.
Legal Entity	An AI system is not a legal entity and existing regulatory frameworks should clarify this position. Notably, as AI systems are continually being improved and developed, there is a possibility that an AI system may gain some form of human sentience. However, this matter falls outside the scope of this submission.
Privacy	This will not be explored in this submission as it falls within the scope of the *Privacy Act Review report* conducted by the Attorney-General’s Department. However, it is worth mentioning that AI systems may process significant volumes of data and both the developer and operator of AI systems need to be responsible to ensure the use of any AI systems does not infringe or violate any applicable privacy legislation or any person’s rights in respect of privacy. Given the risk and potential harm to individuals which may be proliferated and amplified by an AI system, the onus should be

	on operators and developers of AI systems to proactively identify and mitigate such risks.
Intellectual Property	We note that the AI Working Group of the Intellectual Property Policy Group, under the auspices of IP Australia, is currently being tasked with exploring the intersection of AI and intellectual property which we have had the privilege of being involved from providing a legal perspective. As such, this will not be explored in this submission. However, it is worth mentioning that it will be fundamental to clarify issues such as: 1. AI systems not being able to own intellectual property such as copyrights. 2. Who is responsible if AI systems infringe intellectual property rights? 3. Who will be liable if users use publicly available AI systems which generate materials that infringe intellectual property rights of third parties?
Consumer Protection	We note that the intersection between AI and consumer protection are currently being explored by several agencies such as ACCC, ACMA, eSafety Commissioner and OAIC. As such, this will not be explored in this submission.
Tort	Noting the existence and application of torts at common law, it will be desirable if the AI Legislation addresses common and key issues surrounding AI systems and torts which we consider should be legislated for the avoidance of doubt, such as: 1. Liability: One of the most significant challenges is determining who is liable for harm or damage caused by AI systems. In traditional tort law, liability is typically attributed to the person or entity who was negligent in causing or intentionally caused the harm. However, it can be difficult to assign liability in cases involving AI systems, as these systems may be difficult to understand or control. We suggest a proportionate reduction liability regime similar to the regime in the *Civil Liability Act 2002* (NSW) (and the equivalent in each state and territory) be considered given the existence of AI system-enabled supply chains. For certain offences with serious gravitas (for e.g., knowingly developing an AI algorithm which is

	intended to engineer certain results or cause harm), we suggest introducing strict liability offences for both developers and operators of AI systems. We believe that this penalty is balanced given the potential impact of AI systems. It may be worth considering whether the imposition of liability on a user of an AI system is appropriate to put the onus on users to determine whether or not the user should engage with or use an AI system if the AI system is in potential breach of AI Legislation. 2. Bias: Another challenge is regulating AI bias. AI systems are trained on data, and if that data is biased, the AI system will be biased as well. This is likely to be the case given any data is generally historical data. This can lead to unfair, discriminatory and 'engineered' outcomes, such as AI systems that are more likely to make recommendations that systematically adversely affect groups based on race, religion, gender, sexual orientation or other attributes. It is possible that the AI Legislation can cover this topic of "Tort". However, we have not explored this in this submission.
Sector Specific	As the Discussion Paper identified, there are a number of sector-specific regulations, including those for financial services, therapeutic goods, food and beverages, motor vehicles and airline safety. However, we have not explored these in this submission. The development of AI Legislation must consider the impact of AI systems on any sectors of national security interest. These sectors and relevant regulatory frameworks must be identified in this process and there is a need for sector-specific AI Legislation in respect of those sectors of national security interest.

Table 1: Topics and their corresponding proposed regulatory response

Topic	**Proposed non-regulatory response**
AI Ethics Framework	From our understanding, the AI Ethics Framework has been met with a positive reception in Australia. The framework has been praised for its focus on human-centred values, its emphasis on transparency and accountability, and its flexibility to accommodate the evolving nature and use cases of AI

	technology. The framework is a valuable step towards ensuring that AI technology is used in a safe and responsible way. The AI Ethics Framework is a voluntary tool that does not have the force of law. However, it is still commonly used as a guiding tool by developers and deployers in the AI industry. As such, it will benefit from providing more specific guidance on how to develop and use AI systems in a responsible way, perhaps through case examples. The development of this may well be informed by the AI Legislation. **Benefits:** ● More specific guidance could help to ensure that AI systems are developed and used in a way that is consistent with ethical principles. ● More specific guidance could also help to promote innovation in the field of AI, as businesses and organisations would have a better understanding of the ethical considerations that need to be taken into account when developing and using AI systems. ● Case examples could help to illustrate how the principles of AI ethics can be applied in practice. This could be helpful for businesses and organisations that are new to AI or that are not familiar with the ethical considerations involved. **Impacts:** ● Developing more specific guidance could be a time-consuming process and will involve resources being put into the process. ● There is a risk that the guidance could be too rigid or prescriptive, which could stifle innovation. ● It is also possible that the guidance, if not expressed correctly, could be ambiguous.
Technological Unemployment	The government should consider the following strategies to address technological unemployment arising due to AI development: 1. Invest in education and training: As AI technology continues to develop, there will be a need for workers with new skills and knowledge. Australia should invest in education and training programs that will prepare workers for the jobs of the future. This could include programs in coding, data science, and other technical fields. 2. Identify the jobs/sectors to be redundant: A task force

should be implemented to identify the jobs or sectors or industries that will be detrimentally affected due to increased use of AI systems from an employment perspective and identify transitional career opportunities for those to be affected. This task force should provide tools and education to assist career transitions.

3. Create new jobs: AI technology is also creating new jobs, such as those in the development and maintenance of AI systems. Australia should work to create a favorable environment for the growth of these new industries. This could include providing tax breaks or other incentives to businesses that develop and use AI technology.
4. Provide social safety nets: As AI technology displaces workers, there will be a need for social safety nets and support programs to help those who are affected. This could include unemployment benefits, job retraining programs, and other forms of assistance.
5. Collaborate with businesses: Work with businesses and stakeholders to develop a plan for the transition to a more AI-powered economy. This plan should identify the jobs that are most likely to be displaced by AI, and it should outline strategies for retraining workers and creating new jobs.

Invest in education and training

Benefits:

- This would help to ensure that workers have the skills they need to succeed in the new economy.
- It would also help to create a more skilled workforce, which could boost Australia's economy.
- This could help to mitigate the negative effects of technological unemployment.

Impacts:

- It could be expensive to invest in education and training programs.
- It could take time for workers to acquire the new skills they need.
- There is a risk that some workers may not be able to retrain, which could lead to increased unemployment.

Identify jobs / sectors to be redundant

Benefits:

	● This will assist with a targeted approach to the jobs and sectors at risk and which are in need of assistance. ● This will assist facilitate the other strategies outlined. **Impacts:** ● Jobs and sectors identified by the task force will allow affected businesses and individuals to prepare for a career transitional period. ● This will allow the government and businesses to allocate funds accordingly for any transitions. **<u>Create new jobs</u>** **Benefits:** ● This would help to offset the jobs that are lost due to AI technology. ● It would also help to create a more dynamic and innovative economy. ● This could help to boost Australia's economy and create new opportunities for workers. **Impacts:** ● It could be difficult to create new jobs that are as well-paying as the jobs that are lost. ● It could take time for new industries to grow and create jobs. **<u>Provide social safety nets</u>** **Benefits:** ● It would help to mitigate the negative effects of technological unemployment on the economy. **Impacts:** ● It could be expensive to provide social safety nets. ● It could create a disincentive for workers to retrain and find new jobs. ● There is a risk that some workers may become dependent on social safety nets, which could lead to long-term problems. **<u>Collaborate with businesses</u>** **Benefits:** ● This would help to ensure that the government's policies are aligned with the needs of businesses. ● It would also help to create a more efficient and

	effective way to address the challenges of technological unemployment. ● This could help to ensure that Australia is well-positioned to take advantage of the opportunities presented by AI technology. **Impacts:** ● It could be difficult to get businesses to agree on a plan for addressing technological unemployment. ● There is a risk that businesses may not be willing to invest in the skills of their workers.
Government Use of AI	Government use of AI should be encouraged, provided that concerns of security, fairness, accountability and transparency are addressed. This is because AI systems can greatly enhance productivity and capability. We note that the use of AI in government is currently being explored by several agencies such as the Department of Finance, the DTA and the Commonwealth Ombudsman. As such, this will not be explored in this submission.

Table 2: Topics and their corresponding proposed non-regulatory response

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

Answer

For the purposes of this Question 1, we will be only commenting on the following definitions:

- Artificial Intelligence
- Developer
- Deployer
- User

<u>Artificial Intelligence</u>

The Discussion Paper defines **Artificial intelligence (AI)** as "*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 …*".

In our opinion, the following elements of the definition of "Artificial Intelligence" or "AI" is problematic:

- "engineered": We query the need for AI systems to be "engineered". Further, what would constitute an "engineered" system? In our view, it is sufficient to delete the word "engineered" from the definition.
- "predictive": The requirement that an AI system generates "predictive" outputs

adds ambiguity. What would constitute a "predictive" output? Further, in time, it is reasonable to assume that AI systems will have an almost infinite (if not infinite) range of possible outputs. Therefore, putting any restraints on the character of an AI system's output will be counterproductive. In our view, it is sufficient to delete the word "predictive" from the definition.

- "without explicit programming": This requirement is confusing, since all AI systems, to be developed, require some form of programming. In our view, it is sufficient to delete the words "without explicit programming" from the definition.

Altogether, we believe the following refined definition of **Artificial intelligence (AI)** would be more suitable for the AI Legislation: "*a system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives or parameters, regardless of its level of automation*".

In our view, the above refined definition of AI is broad enough to encompass a wide range of AI systems, while also being specific enough to be useful for regulatory purposes. The definition includes the key elements of AI, such as the ability to learn, adapt, make decisions and even create. It also specifies that AI systems must be designed for a specific purpose, which helps to distinguish them from other types of computer systems.

Further consultations with stakeholders will be required for this definition.

Developer

Please refer to Question 20 as to why this definition will be relevant.

Although further consultations with stakeholders will be required, we believe the following refined definition of **Developer** would be a suitable starting position for the AI Legislation: "*a person (and that person's personnel) that designs, codes, produces or develops an artificial intelligence system, whether for its own use or for use by a third party*".

This definition of "developer" should cover different roles that are involved in the development of AI systems, such as:

- Data scientists who collect and prepare data for training AI models
- Machine learning engineers who design and build AI models
- Software engineers who develop the software that uses AI models
- Product managers who define the features and functionality of AI-powered products

The definition would be specific enough to be useful for regulatory purposes, as it identifies the specific activities that are considered to be development of AI systems. This helps to ensure that AI systems are subject to the same regulatory requirements, regardless of who is involved in their development.

Deployer

Please refer to Question 20 as to why this definition will be relevant.

Although further consultations with stakeholders will be required, we believe the following refined definition of **Deployer** would be a suitable starting position for the AI Legislation: "*a person (and that person's personnel) who makes available an artificial intelligence system for users, whether or not for valuable consideration*".

The definition would have the intended effect of extending the AI Legislation to anyone who makes an AI system available for "users" (see below definition).

<u>User</u>

Please refer to Question 20 as to why this definition will be relevant.

Although further consultations with stakeholders will be required, we believe the following refined definition of **User** would be a suitable starting position for the AI Legislation: "*a person who uses an artificial intelligence system*, but excludes a developer and a deployer".

The definition would need to ensure the following:

- The purposes of using the relevant AI system will be irrelevant. This is because the purpose of using it will be irrelevant for a risk-based approach, where it is the risks of the AI system that are examined.
- It will need to extend to the use of AI systems by agents, intermediaries, employees, contractors or others by way of direction. That way, a person or entity will not be able to avoid being classified as a User by mere technicality.

Question 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?

Answer

In our opinion, the following potential risks from AI are not currently covered by Australia's existing regulatory frameworks:

- **Bias and Discrimination:** AI systems can reflect and amplify societal biases, leading to unfair discrimination.
- **Privacy and Data Protection:** While Australia has robust privacy laws, specific provisions for AI applications may be necessary.
- **Deep fakes and Misinformation:** AI-generated deep fakes, which are highly realistic manipulated media, and the spread of AI-generated misinformation are emerging concerns.
- **National Security and Cybersecurity:** AI can have implications for national security, including risks of cyberattacks and adversarial AI.
- **Autonomous Systems:** As AI progresses, autonomous systems, such as self-driving cars and drones, pose unique risks. Australia's existing regulations may need to be updated to address the safety, liability, and accountability challenges posed by these systems. Specific guidelines and standards for testing, validation, and certification of autonomous systems could be developed to ensure their safe deployment.

- **Ethical Use of AI:** While Australia has the AI Ethics Framework, it does not have the force of law.

Possible regulatory actions to mitigate these risks are explored in Tables 1 and 2 in the "Overview" above.

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.

Answer

These are explored in Table 2 in the "Overview" above.

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.

Answer

As explored in Tables 1 and 2 in the "Overview" above, the development and uptake of AI in Australia crosses over multiple different "topics". Any regulatory framework over AI across those different "topics" will need to be unified in their language and approach.

Here are some suggestions on coordination of AI governance across government, along with the goals those suggestions could achieve:

Suggestion	Goals
Establish a central coordinating body	The Australian government could establish a central coordinating body to oversee the development and implementation of AI governance policies and frameworks, including the AI Legislation and any updates to the AI Ethics Framework. This body could be responsible for bringing together different government agencies, stakeholders, and experts to develop a coordinated approach to AI governance. The government should undertake a feasibility study to determine if this body should also serve as the enforcement body under the AI Legislation.
Select a central AI policy framework	The Australian government should select a central AI policy framework that sets out the government's overall approach to AI governance. This framework could include principles and guidelines on how to develop and use AI in a way that is ethical, responsible, and beneficial to society. Given its rather accepted state, the AI Ethics Framework is

	the prime candidate for this central AI policy framework. If the AI Legislation is passed, it is crucial that the central AI policy framework and the AI Legislation “speaks the same language”.
Consult specific sectors on AI governance	The Australian government should also consult the different sectors that will be significantly affected by AI governance frameworks, such as healthcare, finance, and education. These frameworks could tailor the government's approach to AI governance to the specific needs of each sector. This consultation process can be conducted by the central coordinating body mentioned above.

Table 3: Suggestions for coordination across government and their corresponding goals

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?

Answer

The proposed regulations in Tables 1 and 2 in the “Overview” above are similar to those being used or considered for use in the EU and Canada. However, adaptations will need to be made so that a risk-based approach fits the Australian context and the existing frameworks. The reasons are set out in the “Overview” above.

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

Answer

In our opinion, the same approach should apply to public and private sector use of AI technologies. There are various reasons for this.

Firstly, it is difficult to distinguish between public and private sector use of AI, as many AI systems are used by both the public and private sectors. Further, whether used in the public or private sector, the concerns of AI technologies remain largely the same. It would be both efficient and desirable to apply the same approach. A consistent approach will also promote innovation as AI developers will not need to dedicate resources to develop an AI system that could only be deployed in either the public or the private sector. A consistent approach across the public and private sectors also will allow all stakeholders (be it businesses, advisors, customers or government) to develop familiarity with their respective rights and obligations stemming from any AI regulatory framework.

Secondly, while the use of any technology in the public sector will attract additional concerns over security, fairness, accountability and transparency, much like the use of other technologies in the public sector (e.g. emails, personal laptops, data processing software, etc), these concerns can be addressed via governmental or departmental guidelines or internal policies.

Thirdly, Australia would also be aligning its regulatory framework with standards and best practices of other countries in the same WEIRD framework category. This would help to ensure that Australia's regulatory framework is seen as credible and legitimate by businesses and other stakeholders. This will also help promote ease of business for AI business which, given the nature of AI systems, are likely to operate internationally. Similar approaches have been take in the EU and Canada:

1. In the EU, the Artificial Intelligence Act (**AIA**) applies to both public and private sector use of AI. The AIA sets out a number of requirements for the development and use of AI, including transparency, accountability, fairness, and non-discrimination. The AIA also includes specific provisions for the use of AI in sensitive areas, such as law enforcement and healthcare.
2. In Canada, the Standards Council of Canada (**SCC**) has developed a set of AI ethics guidelines that are applicable to both public and private sector use of AI. The SCC guidelines provide guidance on how to develop and use AI in a way that is ethical, responsible, and beneficial to society.

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

Answer

Government use of AI should be encouraged, provided that concerns of security, fairness, accountability and transparency are addressed. This is because AI systems can greatly enhance productivity and capability. Subject to the exploration of the use of AI in government by agencies such as the Department of Finance, the DTA and the Commonwealth Ombudsman, here are some high-level ways that the Australian Government can further support responsible AI practices in its own agencies:

1. Develop and implement a comprehensive AI ethics framework: The Australian Government can develop and implement a comprehensive AI ethics framework that sets out principles for the development and use of AI in a responsible and ethical manner. This framework should be based on international standards and best practices, and it should be regularly reviewed and updated as AI technology continues to develop. Given its rather accepted state, the AI Ethics Framework is the prime candidate for this framework.
2. Create a central coordinating body: See Question 4 above.
3. Provide training and education on AI ethics: The Australian Government can provide training and education on AI ethics to government employees who are involved in the development or use of AI. This training should help employees to understand the ethical implications of AI, and it should help them to develop the skills and knowledge necessary to use AI in a responsible and ethical manner.
4. Establish a process for reviewing and assessing AI systems: The Australian Government can establish a process for reviewing and assessing AI systems to ensure that they are being developed and used in a responsible and ethical manner. This process could include a number of steps, such as:
    a. reviewing the AI system's design and development process;
    b. evaluating the AI system's performance; and
    c. monitoring the AI system's use.

Question 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.

Answer

Generic solutions to the risks of AI are most valuable when the risks are common across a wide range of AI systems. For example, the risk of bias can be a problem for any AI system that is trained on a dataset that is not representative of the population. Generic solutions to this risk can include techniques for debiasing data sets and for ensuring that AI systems are transparent and accountable.

Technology-specific solutions are better when the risks are specific to a particular type of AI system. For example, the risk of privacy violations is a particular concern for AI systems that collect and use personal data. Technology-specific solutions to this risk can include techniques for anonymising data and for ensuring that users have control over their data.

However, we would caution against any legislation mandating technology-specific solutions, given the rapid rate of technology development. There is a risk that technology-specific solutions mandated by regulation will quickly become outdated.

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.

Answer

In relation to Question 9(b), further consultations will need to be had with industry participants. However, here are some critical points in the AI lifecycle where transparency will be valuable:

- **Data Collection and Use:** Transparency is crucial at the data collection stage to inform users about the types of data being collected, the purpose of data usage, and the potential implications for privacy and security. Providing clear explanations and obtaining informed consent can help users make informed decisions about sharing their data.
- **Algorithm Development:** Transparency in the algorithm development phase is essential to understand how AI systems work and to detect and address biases, fairness issues, and potential risks. Providing access to information about the algorithms' design, training data, and decision-making process enables external audits and reviews to ensure accountability and fairness.
- **Model Deployment:** When AI systems are deployed in real-world scenarios, transparency about the model's behavior, decision-making process, and any limitations or uncertainties is critical. Users should have a clear understanding of how the AI system arrived at a particular decision or recommendation.
- **Impact Assessment:** Regular impact assessments throughout the AI system's deployment help identify and mitigate potential risks and ensure ongoing compliance with ethical and regulatory standards. Transparent reporting of these assessments can build public trust by demonstrating responsible AI use.
- **Post-Deployment Monitoring:** Continuously monitoring AI systems in real-world

applications helps identify any unintended consequences or evolving risks. Transparency in sharing monitoring results and any corrective actions taken ensures accountability and fosters public confidence.

Care will need to be taken so that any requirements mandated by regulation should not serve to be a disproportionate barrier to entry for businesses.

In relation to Question 9(b), this is explored in the "Overview" above.

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?

Answer

This is explored in the "Overview" above.

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

Answer

There are a number of initiatives or government actions that can increase public trust in AI deployment to encourage more people to use AI. These include:

1. Take a stance on AI regulations: The government can enact legislation regulating AI in Australia (e.g. the AI Legislation) and embrace the use of AI within government. This will create the impression that use of AI is the norm, rather than the exception.
2. Providing clear and transparent information about how AI systems work: This includes explaining the data that AI systems are trained on, how the systems are making decisions, and how the systems are being monitored and evaluated.
3. Supporting research into AI safety: This research could focus on developing techniques for preventing AI systems from causing harm, such as techniques for ensuring that AI systems are not biased or discriminatory.
4. Enacting regulations that promote responsible AI development and use: Regulations can help to ensure that AI systems are developed and used in a way that is ethical and responsible.
5. Educating the public about AI: This includes teaching people about the potential benefits and risks of AI, as well as how to use AI safely and responsibly.
6. Building public confidence in AI: This can be done by demonstrating the benefits of AI in real-world applications and by ensuring that AI systems are used in a way that is fair and equitable.

Question 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?

Answer

Banning high-risk activities (like social scoring or facial recognition technology in certain circumstances) could have both positive and negative impacts on Australia's tech sector and Australia's trade and exports with other countries.

Positive impacts:

1. Increased public trust and use in AI: Banning high-risk activities could help to increase public trust in AI by demonstrating that the government is taking steps to protect people from the potential harms of AI. This could boost adoption of AI-powered products and services in Australia.
2. More businesses developing and using AI in a responsible and ethical manner: Banning high-risk activities could encourage businesses to develop and use AI in a responsible and ethical manner. This could make Australia a more attractive destination for businesses that are developing and using AI.
3. Australia becoming a more attractive destination for businesses that are developing and using AI: Banning high-risk activities could make Australia a more attractive destination for businesses that are developing and using AI. This is because it would signal to businesses that the government is taking steps to protect people from the potential harms of AI. This will be amplified if public trust and public use of AI has increased.

Negative impacts:

1. Australia becoming less competitive in the global AI market: Banning high-risk activities could make Australia less competitive in the global AI market. This is because other countries, such as Singapore and China, are not as restrictive in their approach to AI development and use. This could give businesses in these countries an advantage in developing and deploying AI-powered products and services. From a policy consideration standpoint, this may be an acceptable outcome if the only innovation stifled are those AI uses with “Unacceptable” risk.
2. Potential for businesses to move their operations to countries with less restrictive AI regulations: Banning high-risk activities could lead to businesses moving their operations to countries with less restrictive AI regulations. This is because businesses may be reluctant to operate in Australia if they are not allowed to use AI in the same way as they can in other countries. From a policy consideration standpoint, this may be an acceptable outcome if the only innovation stifled are those AI uses with “Unacceptable” risk.
3. More difficult for Australian businesses to export AI-powered products and services: Banning high-risk activities could make it more difficult for Australian businesses to export AI-powered products and services to countries that do not have similar restrictions. This is because businesses in these countries may be reluctant to import AI-powered products and services from Australia if they are not allowed to use them in the same way.

In our view, the proposed regulatory framework explored in Tables 1 and 2 in the “Overview” above strikes a proportionate balance between the positive and the negative impacts.

Question 13: What changes (if any) to Australian conformity infrastructure might be required to support assurance processes to mitigate against potential AI risks?	
Answer	This is explored in the "Overview" above.

Question 14: Do you support a risk-based approach for addressing potential AI risks? If not, is there a better approach?	
Answer	For the reasons set out in the "Overview" above, we support a risk-based approach for addressing potential AI risks.

Question 15: What do you see as the main benefits or limitations of a risk-based approach? How can any limitations be overcome?	
Answer	This is explored in the "Overview" above.

Question 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?	
Answer	Whether a particular approach (risk-based or otherwise) is better suited for some sectors, AI applications or organisations than others based on organisation size, AI maturity and resources, will depend on *how* that approach is implemented and enforced. In our opinion, the proposed regulatory framework explored in Tables 1 and 2 in the "Overview" above addresses AI risks proportionately amongst the various sectors, AI applications or organisations. This is because: 1. the proposed regulatory framework targets risk levels posed by the relevant AI system; and 2. while compliance costs will increase with the risk levels posed by the relevant AI system, where the risk levels posed by the relevant AI system is on the higher end, then regardless of the sector, AI application or organisation, the higher compliance costs is arguably justified as the relevant AI system poses higher risks and higher potentials for harm.

Question 17: What elements should be in a risk-based approach for addressing potential AI risks? Do you support the elements presented in Attachment C?	
Answer	This is explored in the "Overview" above.

Question 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?	
Answer	As explored in the "Overview" above, any successful and robust regulation of AI will

involve many different legal “topics”. In Australia, these topics are governed by different existing frameworks.

While a risk-based approach, such as the AI Legislation, is desirable, it will need to be accompanied by other regulatory changes. These have been set out in Tables 1 and 2 in the “Overview” above.

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

Answer

A risk-based regulatory approach, such as the AI Legislation, is flexible and adaptable in its regulation of AI. By focusing on the risks posed by an AI system (i.e. a somewhat functional assessment of the AI system), the AI Legislation would be well equipped to account for the various use cases of AI systems (e.g. LLMs and MFMs), both now and in the future.

Question 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?

Answer

In relation to Question 20(a), please see our response to Question 6.

In relation to Question 20(b), in our opinion, a risk-based approach will only be truly effective if it applies across the board - i.e. to developers, deployers and users of AI systems.

- Developers
  - **Developers are responsible for the design and development of AI systems.** They have the most control over the features and capabilities of these systems, and they are therefore in the best position to identify and mitigate risks.
  - **Developers are best placed to conduct rigorous testing.** Rigorous testing is most suitably done before an AI system is fully developed and deployed. At this stage, the AI system remains under the control and development of the developer.
- Deployers
  - **Deployers are responsible for the deployment of AI systems.** Deployers are responsible for the deployment of AI systems, and therefore have significant control over the risks associated with those systems entering the real world.
  - **Deployers have a duty of care to the users of their AI systems.** They must ensure that these systems are used in a safe and responsible manner, before deploying the AI system out to users.
  - **Deployers often collaborate with developers.** AI systems are often the product of a collaboration between developers and deployers. It would

make sense to regulate deployers alongside developers.

- Users
    - **Users are responsible for their own actions when using AI systems.** Much like the internet, AI systems are neutral but powerful information repositories. Some AI systems can also act on that information. Users should remain accountable for what they do with those AI systems.

Annexure B – Proposed Risk-based Approach for Australia

Risk Profile	AI System	Regulatory Framework
Low	● AI-enabled video and computer games ● Spam filters ● AI systems that offer product recommendations or personalised content ● AI-based language translation services ● AI systems that can autonomously (without human intervention) create a separate AI system that carries the above risks	Developer obligations: ● None Deployer obligations: ● None User obligations: ● None
Medium	● Human impersonation (i.e. chatbots) ● AI applications used in hiring and employee evaluation ● AI systems used for crime prediction and law enforcement ● AI used for content moderation on social media platforms ● AI systems that can autonomously (without human intervention) create a separate AI system that carries the above risks	Developer obligations: ● Notify humans that they are interacting with an AI system ● Notify humans that emotional recognition or biometric categorisation system are applied to them ● Notify humans of the basis of the AI's decision making ● Apply labels to deep fakes Deployer obligations: ● Ensure that the AI system complies with the Developer obligations before deploying the AI system

		User obligations: ● None
High	● AI used for ‘real-time’ and ‘post’ remote biometric identification of people ● AI used as safety components in the management and operation of road traffic and the supply of water, gas, heating and electricity ● AI that is used in the education sector and to determine further access to education ● AI used for recruitment or evaluating job candidates, or for monitoring and evaluation of employees ● AI used by public authorities that determines access to public assistance benefits and services ● AI used to evaluate the creditworthiness of people (with the exception of AI systems put into service by small scale providers for their own use) ● AI to dispatch, or to establish priority in the dispatching of emergency first response services, including by firefighters and medical aid ● AI that assesses the risk of offending, reoffending, victimhood ● AI used by law enforcement to detect people’s emotional states (e.g. as polygraphs) ● AI used by law enforcement to detect deep fakes ● AI used by law enforcement in connection with a criminal offence (evaluating the reliability of evidence, predicting the occurrence or reoccurrence of an actual or potential criminal offence, profiling people, identifying patterns, assessing risk) ● AI systems used by public authorities to assess a risk ● AI used by public authorities to verify documents ● AI that assists public authorities in examining applications for asylum, visa and residence permits and associated complaints ● AI used for critical medical diagnoses or treatment recommendations ● Self-maneuvering vehicles ● AI systems that provide financial services or that make high-stakes financial decisions	Developer obligations: ● Use high-quality training, validation and testing data ● Establish documentation and design logging features ● Ensure appropriate degree of transparency ● Ensure robustness, accuracy and cybersecurity Deployer obligations: ● Establish and implement quality management ● Keep up-to-date technical documentation ● Undergo compliance assessment and reassessment (for modifications) ● Display signs or declarations of compliance ● Conduct routine monitoring ● Collaborate with surveillance authorities User obligations:

	● AI systems that can autonomously (without human intervention) create a separate AI system that carries the above risks	● Operate AI system in accordance with instructions of use ● Ensure human oversight ● Monitor for possible risks ● Cease using the AI system if they are aware, or if a reasonable user would be aware, that the use of that AI system will reasonably likely materialise the risks
Unacceptable	● AI-powered weapons systems that can autonomously (without human intervention) make or carry out decisions that harm life, limb or property ● AI systems designed for cyberattacks, hacking, or other harmful activities ● AI used for creating and distributing non-consensual explicit content (deep fakes), child exploitation or human trafficking ● Practices that have a significant potential to manipulate persons through subliminal techniques ● Practices that exploit the vulnerabilities of specific vulnerable groups (e.g. children, persons with disabilities) ● AI-based social scoring done by public and private authorities ● AI systems that can autonomously (without human intervention) create a separate AI system that carries the above risks	Developer obligations: ● Development of such AI systems is banned Deployer obligations: ● Deployment of such AI systems is banned User obligations: ● Use of such AI systems is banned

General Comments:

- You will note that, for reasons set out in the "Overview" above, the table above was largely based on the AIA in the EU. However, changes have been made to account for the Australian context. The elements of Attachment C of the Discussion Paper have also been considered.
- It will be important to ensure that the AI systems for each risk profile are non-exhaustive. This is because AI technology develops rapidly, and an exhaustive list may be very short lived.
- Testing or beta stages:

    - A developer should not be considered to have breached its obligations if the AI system is still in a testing or beta stage. Like the development of "traditional" software, the software may have bugs or glitches that achieve unintended consequences. Developers should be given reasonable opportunity to address those unintended consequences and comply with their obligations.
    - However, care should be taken in the drafting so that a developer cannot claim that an AI system is still in a testing or beta stage, when it is in fact fully developed and made available to deployers and/or users.
- Regulatory sandbox:
    - As further set out in the "Overview" above, some allowance should be given to developers and deployers who operate within any regulatory sandbox that is set up.
    - However, AI systems with an "Unacceptable" risk profile, due to their nature, should remain prohibited.
- The proposed AI Legislation should not impose disproportionate burdens to developers, deployers or users. Exceptions should be considered for AI systems with Medium or High risk profiles, whereby developers, deployers or users will only need to implement controls to a level reasonably required to manage those risks. This is not an uncommon position taken by regulation. For example, in *Australian Securities and Investments Commission v RI Advice Group Pty Ltd* [2022] FCA 496, the Federal Court clarified, amongst other things, that holders of Australian Financial Services licences must "materially reduce cybersecurity risk through adequate cybersecurity documentation and controls to an acceptable level" by implementing controls to a level reasonably required to manage those risks.