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

The definition of AI needs to be expanded beyond the current narrow definition, which states that AI "**generates predictive outputs without explicit programming"** for a given set of objectives or parameters. This definition fails to encompass the capabilities of different AI solutions today, particularly Generative AI in creative domains.

Generative AI has demonstrated outputs that surpass the limits of predictability by the human mind. For instance, when prompted to generate an image of a young woman walking in a business suit on a rope between two skyscrapers with sharks in the water below, and the ability to see the Sydney Opera House on one side and an Arabian desert with camels on the other, along with several other world wonders in the sky, AI can produce such creative and imaginative results.

It is important to note that any AI solution comprises a user interface and backend computation. Claiming that AI makes decisions without programming is inaccurate. Regardless of whether the prompts are in plain English or complex queries, AI involves advanced programming and computation to function effectively.

Suggested definitions :
AI - Artificial Intelligence (AI) refers to a collection of capabilities powered by technology, data and advanced computation which performs tasks and delivers outputs similar to or better than human intelligence in certain areas.
Other definitions for consideration
AI solution - Product or service which uses in part or in whole AI related components to deliver outputs to customers (external or internal).
Critical service -
Those critical elements of infrastructure (namely assets, facilities, systems, networks or processes and the essential workers that operate and facilitate them), the loss or compromise of which could result in:
a) Major detrimental impact on the availability, integrity or delivery of essential services - including those services whose integrity, if compromised, could result in significant loss to finanical, health & safety, loss of life.
b) Significant impact on national security, national defence, or the functioning of the state.
(note - uses certain language from
https://www.npsa.gov.uk/critical-national-infrastructure-0)
https://www.homeaffairs.gov.au/nat-security/files/cic-factsheet-what-is-critical-infrastructure-centre.pdf

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?

Government should model and provide guidance on response and recovery from unintended consequences of AI solutions, including crisis management. Lessons learnt from the major security events in large private enterprises should provide as a case study on the impacts to customers and general public.
Example: AI solutions involved in airlines, clinical research & pattern recognitions.

Potential risks:
1. No established mechanisms for customers/ agencies/ institutions receiving service to genuinely challenge the outcomes of service(e.g. bias expected).
2.No processes or practices established to manage any outage/crisis/ deficiency in delivery of services from AI solutions. Can include absence of roll back, fall back, alternate options with reduced service performance etc.
3.No scenario modelling or simulation done at regular intervals to test the usage of AI solutions for critical services. (i.e. this is mimicking more real time scenario than traditional testing approaches before release of AI solutions). Example- AI looks at traffic patterns and closes lane in Harbour bridge and alerts more buses in X route, when done incorrectly leading to overall chaos in Sydney traffic.

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.

Government should proactively educate the small and medium business and public on the potential and risks of AI. Leaving small and medium business with no ability to compete with big global and Australia corporations will be detrimental to the economy.
E.g. small design and marketing agency will be under threat of Gen AI self-serve services and equally a local training provider services on health and safety could be taken over by far better AI simulation content and trainers.
While disruption is unavoidable, giving fair opportunity for businesses to adapt to the changes is essential.

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.

Government should set up a body for overseeing and administering AI Assurance- for all government agencies, for critical services in large enterprises using AI solutions.
The intent is not to overregulate but to foster active, fair, and safe use of AI for the benefit of customers and citizens. It is crucial to convey the correct narrative regarding AI assurance from the government.

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?

Australia could look into UK's leadership in setting up portfolio of AI Assurance techniques. While we don't recommend government to take on a licensing role of AI systems, it should play a Lead Assurance Partner role to set guardrails and control mechanisms.
https://www.gov.uk/guidance/cdei-portfolio-of-ai-assurance-techniques

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

The controls for public and private sector enterprises should be common for 'critical services' using AI solutions. The level of transparency and assurance on usage of AI in public services should be monitored and governed with more rigour, while it will be incumbent on the executives of private

sector to do the same. In connected world, any impact to critical service irrespective of it being public or private could be severe and will need adequate governance oversight without needing to overregulate it.

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

Government should roll out a controls framework that provides a comprehensive view of the controls to be applied across the lifecycle of AI solution.
Resonvate' s AI Controls Framework (AICF) has been evolved and adopted in products and solutions using AI.
Resonvate is open to share its controls framework for the Australian Government to evolve its thinking on AI controls.
RAICF covers 5 domains and 22 control areas - Transparency, Communicating on AI, Monitoring & Assurance, Accountability, AI Design and Build

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.

To make risk management effective,
generic solutions for managing risks are valuable, while handling matters related to topics such as skills, accountability, assurance, governance, third party services, privacy, finance, safety, marketing of products & services, compliance to regulations, AI lifecycle processes etc.
technology specific solutions for managing risks are valuable, while dealing with the below the radar items of all things AI.
few examples would be-
a) algorithmic decisions related to usage of techniques like linear regression, supersparse linear integer model for scoring, known neighbour network for recommendations etc.., as these might have impact on the outputs
b) data decisions related to quality and quantity of data to be used for training, decision making & frequency of refresh;
e.g. correlating citizen's behaviour while paying fines in tax, crime, parking, registration, licenses etc. to recommend policy decisions
c) architecture decisions related to the technology solution components, APIs and processes use for running and processing the AI models
e.g. security protocols and API authentication used by a tax defaulter prediction AI using Medicare, insurance, and health data
d) fairness decisions to ensure models have not explicitly or implicitly introduced bias while making decisions
e.g. person from particular country, religion, occupation blacklisted or given lesser preference while inclusion for eligibility of a government benefits scheme, due to bias created from generic perceptions of the data cluster

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?

Actions proposed arising from AI models need validating against AI algorithms such that a desk-based calculation undertaken manually during the AI development phase so that they can repeat the analysis to conclude with the same action that can be confirmed by the sponsors of the AI initiative. Supports the key AI requirement that all AI decisions/actions can be supported by a human. We need to eliminate the justification that 'The system says that'. Actions proposed from AI cannot be linear without boundaries else they lead to actions unforeseen by project sponsors.

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

Governance of AI is no different than governance over other requirements other than the potential lack of appreciation by project sponsors that the model could lead to unforeseen actions. This would mean that testing of AI would need to test all range of actions that could result to confirm model boundaries. Sponsors would be accountable for AI actions much as they are accountable today for all project requirements. Projects will need to undertake mandatory testing of the model to confirm boundaries of AI actions. Depending on the risk profile of the AI actions these test results will need to be agreed by sponsors and in some cases escalated to higher governance based on the risk assessment of the AI actions that can be taken. For example if AI actions could lead to very impactful harm to a 3rd party (customer for example) then this would mandate greater governance oversight and accountability for confirmation of the AI model requirements.

10. Do you have suggestions for:

a. Whether any high-risk AI applications or technologies should be banned completely?

This is simple: assessment of the AI actions and impacts is similar to how organisations apply delegated authorities. AI policies need to align to accountabilities in the organisation such that someone is responsible for the agreement of the proposed action envisaged in the AI model. This may mean agreement as to the extent of real-time oversight needed to confirm specific AI actions. If there isn't a practical individual to confirm the criteria for the AI actions or implement real-time oversight then that would indicate that the use of AI is inappropriate as the fundamental rule is that a human must be accountable for all AI actions/outcomes and that this individual must be appropriately qualified and who can explain the rationale behind the AI action.

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

As above when there is no individual qualified to take accountability for the AI actions then AI cannot be applied.

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

By implementing in law/regulations the fundamental rule of human accountability for any actions. That any use of AI will need to show transparency in identifying an individual, appropriately qualified who can justify and if necessary, approve AI actions. Every use of AI will be listed in a transparent register with outcome accountability linked to job role. Annual audits will confirm use of AI register and appropriateness of accountabilities.

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?

Refer response for #11

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

AI usage should be incorporated into annual Audit requirements. Small companies using AI will need to register AI usage and maintain accountability registers.

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

Yes, risk-based approach is a practical way to manage AI solutions and adoption. However, it needs to be mapped to organisational governance roles rather than high, medium, low approach - need to align risk to responsibility else it is too esoteric. Aligning AI related risks to existing risk management frameworks will be critical.
Non-financial organisation can adopt leading practices from APRA's Banking Executive Accountability Regime (BEAR) regime which clearly expects the senior management and executives to be accountable for managing risks.
Reference: https://www.apra.gov.au/sites/default/files/2020-12/BEAR%20information%20paper%20December%202020.pdf

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

Risk approaches without accountabilities just create a number, How is this used? Personal accountability alignment is needed.

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?

Also equally relevant is whether individuals in the organisation are competent to perform their role as being accountable for AI outcomes. Unless you can find someone who is competent then you cannot use them as to be accountable for AI outcomes. This would infer that a degree of AI capability building is needed, much as we think of people needing to build their competencies in a range of areas. Whether this is publicly certifiable or not it comes down the organisation being comfortable that an individual in a role is competent in AI governance.
There should not be any variations in approach to sectors.
The variations can be applied while determining AI solution is applied for a critical or non- critical service,
(Please refer response to #1 on definition of critical services)

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

By applying Boundaryless Thinking (borrowing ideas from cross-industry domains), Australian government can provide risk guidelines for different degree of autonomy delivered in AI solutions. There should be consideration to build a risk framework for AI similar to the level of autonomy concepts used in autonomous vehicles.
Society of Automotive Engineers have defined level of autonomy standard and have been improving it for many years now.
https://www.sae.org/standards/content/j3016_202104/

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?

AI though has its nuances and depth while dealing with related risks, it is better off to be integrated into existing risk management frameworks and practices of government and private sector.
The existing risk frameworks should be strengthened to ensure -
- AI related risks are adequately identified
- Risk treatment plans are specific to AI
- Risk monitoring and governance is equally applied on AI
- Data & reporting on AI risks are made available on demand

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

The key is to set up guardrails and assurance in a 3-plane approach.
1. DATA PLANE - holds the meta data, data, training data, logic, algorithms (e.g. Da-Vinci model)
2. USER PLANE - helps to configure the use cases for AI solutions (e.g. use for medical record triage, prompts to suggest tax regulations etc.)
3. CONTROL PLANE - Used to define and control the different settings to regulate the user plane
While Data Plane could be owned by certain companies which specialise in data models and provide them for other parties to build on them. This layer will be difficult to regulate and might need to rely on existing regulations or external assurance reported by these firms.
User plane and Control plane can be better regulated with right controls and assurance mechanisms.

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?

Challenge is that if it is applied only to organisations of certain sizes then they will create small companies to get around the requirements. AI can be used by small organisations so all need to be included. An organisation must declare its use of AI much like they do with privacy. Any entity that transacts with Australian citizens - similar as to how EU applies its privacy principles, but how do you regulate overseas use of AI other than to demand an Australian individual is accountable for outcomes and that probably is impossible to mandate.

b. developers or deployers or both?

Developers cannot be held accountable since they are part of the IT industry which is currently unregulated and very difficult to regulate. Better to regulate their customers.