Safe and responsible AI in Australia

Proposals paper for introducing mandatory guardrails for AI in high-risk settings

KPMG Australia, October 2024
KPMG.com.au

Contents

Executive summary

As a leading professional services firm, KPMG Australia (KPMG) is committed to meeting the requirements of all our stakeholders – not only the organisations we audit and advise, but also employees, governments, regulators – and the wider community. We strive to contribute in a positive way to the debate that is shaping the Australian economy and we welcome the opportunity to provide a submission in response to the *Introducing mandatory guardrails for AI in high-risk settings* proposals paper.

KPMG is an early and active user of AI, having partnered with Microsoft to streamline the deployment of AI in our back-office functions and consider its use across tax, audit and advisory work.[1] Through our Trusted AI Framework, KPMG is also committed to a human-centred approach to responsible AI that we apply to the design and deployment of AI within the firm. KPMG is considered both a developer and deployer of AI applications in Australia and has been on a journey over the last few years, evolving a robust AI governance structure and putting in place appropriate policies and processes that underpin our approach to responsible AI. KPMG is also working towards a certification against the ISO42001:2023 AI Management System Standard.

KPMG supports introducing a new cross-economy AI-specific Act that provides clear and consistent expectations on those developing and deploying AI, better enabling interoperability with international approaches, and addressing complexity and duplication within existing legislative frameworks. We note that the EU AI Act was approved by the European Union in March 2024, and many Australian businesses who have cross border clients, suppliers, or data stored in the EU are already required to comply with this regulation. Australia's regulatory framework should aim to align closely with this in order to reduce administrative burden on businesses.

The successful adoption of responsible AI needs to be assisted by addressing the public's current lack of trust in AI by ensuring the right mix of policy settings, regulations and laws to ensure AI use is safe. KPMG supports the Government's commitment to developing a regulatory environment that builds community trust and promotes AI adoption, especially in high-risk settings. In this submission KPMG responds to the consultation questions in the proposals paper on the proposed guardrails, defining high-risk AI, and the regulatory options for mandating the guardrails.

KPMG supports the development of principles that define high-risk AI, supported by examples in a non-exhaustive list which gives industry guidance and clarity on how the principles operate in practice. While the principles in the proposals paper provide a comprehensive starting point, consideration should be given to a number of additional factors to strengthen these, including interconnections between different risks, ensuring sensitive information is appropriately captured, and making specific reference to culturally informed consultation.

The submission outlines 14 recommendations at section one and directly addresses the consultation questions at section two. If you would like to discuss the contents of this submission further please do not hesitate to reach out. KPMG looks forward to continuing engagement with the Australian Government as it develops a safe and responsible framework for AI in Australia.

Yours sincerely,

John Munnelly
Chief Digital Officer
KPMG Australia

James Mabbott
Partner in Charge, KPMG Futures
KPMG Australia

Kelly Henney
National Leader, Privacy & Data Protection
KPMG Australia

[1] KPMG and Microsoft agreement to put AI at the forefront of professional services – Media release 12 July 2023

Background

About KPMG

KPMG Australia is an Australian partnership and a member firm of the KPMG global organisation of independent member firms affiliated with KPMG International Limited (KPMG International), with more than 270,000 partners and employees spanning 140 countries and territories around the world. KPMG Australia makes a strong economic contribution, employing close to 10,000 people and partners across Australia.

KPMG Australia is committed to the responsible development and use of artificial intelligence as articulated in the KPMG Trusted AI Framework and implemented through the firm's AI management system. KPMG is also working towards a certification against the ISO42001:2023 AI Management System Standard.

This submission builds on KPMG's previous engagement in the safe and responsible development of AI in Australia and globally. KPMG has provided a number of submissions to various forums on this topic, including on Safe and Responsible AI in Australia in August 2023, Automated Decision Making and AI regulation in July 2022, An AI Action Plan for all Australians in December 2020, the Australian Data Strategy in July 2022, and Human Rights and Technology in 2020 and Beyond in March 2020. KPMG published a report with the AIIA in March 2023, Navigating AI: analysis and guidance on use and adoption, which examines the global and domestic regulatory landscape in the Artificial Intelligence space. KPMG has published a number of other relevant reports on AI, including *A Prosperous Future: Emerging Tech* in collaboration with AmCham Australia in 2022, *Top risks to Australian Business 2024-25* in 2024, and *AI Amplified: What Gen Zs think of AI* by Year 13 in collaboration with KPMG and Microsoft in 2024.

We have also done extensive work with the University of Queensland on the topic of *Trust in Artificial Intelligence*. The most recent paper, Trust in Artificial Intelligence: Global Insights 2023, was published in February 2023 and surveyed over 17,000 people from 17 countries on the public's trust and attitudes towards AI. Previous work in this series includes Achieving Trustworthy AI: A Model for Trustworthy Artificial Intelligence, Trust in Artificial Intelligence: A five country study, and Trust in Artificial Intelligence: Australian Insights 2020. KPMG Australia is also a proud Anchor Partner of the Human Technology Institute, a cornerstone in our pursuit of Trusted AI.

Section 1:
KPMG recommendations

KPMG recommendations

Recommendation 1: *KPMG supports the development of principles that define high-risk AI, and suggests that the following be considered when finalising the principles:*

- *A comprehensive review of the overall risk environment, so that interconnections between AI and other risk areas can be understood and assessed;*
- *Introduction of a system to test and govern AI and emerging technologies that are imported from other jurisdictions against Australian human rights, data protection and related laws and ethical standards;*
- *Consideration of a principle that captures high-risk AI systems that process sensitive information, given the privacy risk to individuals; and*
- *Consideration of the risk of adverse impacts to global security and cooperation as part of principle (e).*

Recommendation 2: *The principles will need to be applied on a 'use case' basis. Assessing high-risk AI on an application basis will not protect against harm that may arise in different use cases.*

Recommendation 3: *KPMG recommends that the principles be strengthened by making specific reference to culturally informed consultation in adherence to Free, Prior and Informed Consent (FPIC) principles, and by referencing the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP).*

Recommendation 4: *KPMG supports the adoption of a principles-based approach. In order to provide sufficient clarity and certainty, it will be important for the principles to be supported by examples in a non-exhaustive list which gives industry guidance on how the principles operate in practice.*

Recommendation 5: *KPMG considers that there are certain AI applications or Automated Decision Making (ADM) tools that may be damaging and may undermine fundamental values of our society, including democracy and human rights and, therefore, should be banned...*

Recommendation 6: *KPMG considers that the principles are flexible enough to keep up with emerging technologies. At the same time, business commitment to strong AI ethics will help in controlling for emerging AI technologies that cause significant harm. To ensure any mandatory guardrails or regulatory framework keeps up with emerging technologies, we suggest a thorough government led review is undertaken at least every two to three years and that any significant movement in international regulatory frameworks are considered and adopted as soon as practical.*

Recommendation 7: *KPMG considers that given the reach of general-purpose AI systems, mandatory guardrails should apply to these models. It may be useful to consider Section 110 of the EU-AI Act which outlines possible systemic risks of general-purpose AI models.*

Recommendation 8: *KPMG considers that any Australian framework should mirror the existing key indicators for defining high-risk GPAI models from the EU AI Act.*

Recommendation 9: *KPMG considers that the implementation of assurance mechanisms would facilitate greater trust in high-risk AI systems. For guardrail 10 to be effective, it will be important for an effective assessment and assurance framework to be developed to ensure that these assessments are meaningful.*

Recommendation 10: *KPMG recommends that the government consider whether an additional guardrail needs to be added to address the safe decommissioning of high-risk AI systems. Where high-risk AI systems are not safely decommissioned, this could cause harm to any users of these systems including members of vulnerable populations.*

Recommendation 11: *KPMG considers that there could be merit in governments mandating or preferencing suppliers that are accredited to a certain industry standard (for example ISO42001) when procuring AI. This could also be an interim measure before a formal conformity / assurance framework is developed for high-risk AI.*

Recommendation 12: *To reduce the regulatory burden on small-to-medium sized businesses applying guardrails, the Government could consider providing support to businesses to help them prepare for any regulatory changes, including education programs, and leveraging third-party AI providers who can provide protections as part of their services.*

Recommendation 13: *KPMG supports introducing a new cross-economy AI-specific Act that provides clear and consistent expectations on those developing and deploying AI, better enables interoperability with international approaches, and addresses complexity and duplication within existing legislative frameworks.*

Recommendation 14: *Greater consistency with international regulatory frameworks would significantly reduce administrative burden, help with exporting technology out of Australia, and set clearer expectations for the importation of technology.*

Section 2:
KPMG insights

Defining High Risk AI

Consult questions

1. Do the proposed principles adequately capture high-risk AI? Are there any principles we should add or remove?
    - Please identify any: low-risk use cases that are unintentionally captured

KPMG Response Q1

KPMG supports the development of principles that define high-risk AI. The proposed principles provide a comprehensive starting point and would capture a significant amount of high-risk AI. We also welcome their alignment to principles in the EU AI Act, noting that Australian entities that have suppliers and clients in the EU will be captured by this Act, so alignment will be important to ensure consistency and lower the cost of compliance. However, KPMG suggests there are several additional considerations when finalising the principles. Importantly, the principles will need to be applied on a ‘use case’ basis. Assessing high-risk AI on an application basis will not protect against harm that may arise in different use cases. For example, using AI for a training video may be relatively low risk, however, using the same application for political campaigns could be deemed high risk.

The nature of AI requires a comprehensive and systemic understanding of risk dynamics. As explored in KPMG's recent Top Risks to Australian Businesses report, there are mutually reinforcing relationships between AI risks and the rising risks associated with political polarisation and misinformation / disinformation. The risk represented by any one AI technology in isolation may not be deemed high, but consideration would be needed regarding the way that technology enables and magnifies a wide range of other risks. For example, when the risk represented by an AI that creates realistic and convincing fake videos is combined with the growing risk of political polarisation and civil unrest, the potential negative impact becomes greater than the sum of its parts. For this reason, KPMG recommends a comprehensive review of the overall risk environment, so that the interconnections between AI and other risk areas can be understood and assessed. This could be considered under principle (e), *the risk of adverse impacts to the broader Australian economy, society, environment and rule of law.*

Australian organisations deploying AI and emerging technology solutions may face significant challenges with the use or application of technologies developed in jurisdictions with diverging and potentially conflicting human rights standards and protections. To address this, KPMG recommends that policy makers introduce a system to test and govern AI and emerging technologies that are imported from other jurisdictions against Australian human rights, data protection and related laws and ethical standards. This could be based on an internationally endorsed accreditation system developed by a recognised international standards body, where practical.

KPMG also supports the proposal in the Privacy Act Review to regulate activities with high privacy risks which would capture some AI systems and technologies. We note that this is covered in principle (a), *the risk of adverse impacts to an individual’s rights recognised in Australian human rights law without justification, in addition to Australia’s international human rights law obligations*.

KPMG suggests there should be a principle that captures high risk AI systems that process sensitive information, as there is a high risk associated with AI systems intruding on an individual’s privacy without seeking the individual’s express consent. When organisations or agencies use data containing personal information and/or sensitive information to design and operate AI solutions, it is critical that individuals are made aware of what data is being collected and how that will be used. Both organisations and agencies must be required to take precautions to protect privacy and provide individuals with the opportunity to opt out, enabling the individual the right to exercise control over their personal and sensitive information. For example, the EU AI Act has documented AI systems processing biometric data as prohibited use of AI. This is due to AI solutions having the capability to categorise individuals based on their biometric data to deduce or infer an individual's race, political opinion, trade union membership, religious or philosophical beliefs, sex life or sexual orientation. These data attributes are classified as sensitive information within Australia, and similarly within Europe they are classified as special categories of personal information. There are some limited exceptions noted in the EU AI Act in the context of law enforcement, but this will need to be clearly defined and articulated to avoid ambiguity.

In relation to principle (e), *the risk of adverse impacts to the broader Australian economy, society, environment and rule of law*, KPMG suggests that the Government should also consider the risk of adverse impacts to global security and cooperation, as geopolitical competition drives countries to limit international collaboration in favour of closed networks with ideological allies. Importantly, categories of uses that relate to defence or national security should be treated separately.

Recommendation 1: *KPMG supports the development of principles that define high-risk AI, and suggests that the following be considered when finalising the principles:*

- *A comprehensive review of the overall risk environment, so that interconnections between AI and other risk areas can be understood and assessed;*
- *Introduction of a system to test and govern AI and emerging technologies that are imported from other jurisdictions against Australian human rights, data protection and related laws and ethical standards;*
- *Consideration of a principle that captures high-risk AI systems that process sensitive information, given the privacy risk to individuals; and*
- *Consideration of the risk of adverse impacts to global security and cooperation as part of principle (e).*

Recommendation 2: *The principles will need to be applied on a 'use case' basis. Assessing high-risk AI on an application basis will not protect against harm that may arise in different use cases.*

2. Do you have any suggestions for how the principles could better capture harms to First Nations people, communities and Country?

KPMG Response Q2

The rapid and escalating adoption of AI technology represents a unique potential harm to First Nations people, both as users of the technology and rights holders of land adjacent to AI data storage facilities. These potential impacts to First Nations people include:

- **Bias and discrimination:** AI systems can inadvertently perpetuate existing biases and discrimination. This can occur when the data used to train the AI system reflects past human discrimination or when minority populations, such as Aboriginal and Torres Strait Islander people, are insufficiently represented in the data. This can lead to AI systems that do not serve the needs of these populations or actively undermine them.
- **Violation of rights:** The use of AI systems can potentially violate the rights of Aboriginal and Torres Strait Islander people. For example, if AI systems are used to make decisions that affect Aboriginal and Torres Strait Islander people without their Free, Prior, and Informed Consent, this could violate their rights as outlined in the UN Declaration on the Rights of Indigenous Peoples.
- **Cultural impact:** AI systems can potentially have a negative impact on the culture of Aboriginal and Torres Strait Islander people. For example, if AI systems are used to digitise and disseminate cultural knowledge without the consent of Aboriginal and Torres Strait Islander people, this could lead to cultural appropriation and the loss of control over Indigenous cultural and intellectual property.
- **Environmental and health impacts:** The construction and operation of AI data storage facilities on Aboriginal land and accessing Aboriginal waters could lead to environmental degradation, including pollution of land and water resources, disruption of ecosystems, and loss of biodiversity, which would

have profound impacts on the traditional lifestyles and cultural practices of Aboriginal people, as well as health problems related to the consumption of contaminated water or food.

- **Cultural heritage impact:** The construction of AI data storage facilities could potentially lead to the destruction or alteration of sites that are of cultural, spiritual, or historical significance to Aboriginal and Torres Strait Islander people. This could result in the loss of cultural heritage and a disruption of cultural practices and traditions. Furthermore, it could potentially lead to changes in local economies and social structures, which could have negative impacts on Aboriginal and Torres Strait Islander communities, such as increased competition for resources, changes in employment patterns, and increased inequality.

It is crucial that the rights of First Nations people are respected and protected to mitigate against these potential harms, that they are fully consulted and involved in decision-making processes related to the construction and operation of AI data storage facilities, and that mechanisms to ensure First Nations perspectives and protection against potential impacts are embedded into the design of AI systems. This could be achieved by:

- **Inclusive consultation**: Engage First Nations people in the development and implementation of AI systems. This can help ensure that the systems are culturally sensitive and do not inadvertently harm First Nations communities or Country.
- **Cultural awareness:** Incorporate cultural awareness into AI systems. This can help ensure that the systems respect and uphold the rights, traditions, and values of First Nations people.
- **Data sovereignty**: Respect the data sovereignty of First Nations people. This means recognising that First Nations people have the right to control the collection, use, and storage of data that pertains to them.
- **Impact assessments**: Conduct impact assessments to identify and mitigate potential harms to First Nations people, communities, and Country. These assessments should consider, among other issues, the social, cultural, and environmental impacts of the construction of AI data facilities.

To achieve this, the proposed principles for safe and responsible AI in Australia could be strengthened by making specific reference to culturally informed consultation in adherence to Free, Prior and Informed Consent (**FPIC**) principles, and by referencing the United Nations Declaration on the Rights of Indigenous Peoples (**UNDRIP**), which provides a universal framework of minimum standards for recognising and protecting the unique rights of Indigenous people.

While it is critical that appropriate safeguards are in place to avoid potential harm to First Nations people and other underrepresented groups, KPMG also notes the significant value that AI can provide to these groups. A recent *AI Amplified: What Gen Zs think of AI* report, developed by Year 13 in collaboration with KPMG and Microsoft, surveyed Australian youth to understand young people's perceptions, interests, and engagement with AI. A key finding of the report was how gender, socioeconomic status, and culturally and linguistically diverse backgrounds are significant determinants shaping interest and understanding levels towards AI, with males and people from low socioeconomic and culturally and linguistically diverse backgrounds showing heightened engagement with AI compared to the average. This demonstrates the potential for generative AI tools to democratise access to knowledge, skills and information with young people from marginalised backgrounds.

Recommendation 3: *KPMG recommends that the principles be strengthened by making specific reference to culturally informed consultation in adherence to Free, Prior and Informed Consent (FPIC) principles, and by referencing the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP).*

3. Do the proposed principles, supported by examples, give enough clarity and certainty on high-risk AI settings and high-risk AI models? Is a more defined approach, with a list of illustrative uses, needed?
   - If you prefer a list-based approach (similar to the EU and Canada), what use cases should we include? How can this list capture emerging uses of AI?

- If you prefer a principles-based approach, what should we address in guidance to give the greatest clarity?

KPMG Response Q3

KPMG supports the adoption of a principles-based approach with practical requirements that take into consideration principles such as human-centred design, transparency, explainability and interpretability, data minimisation, lawfulness and fairness, purpose limitation, accountability, security, user control and consent, ethical use of data, an individual's right to access, rectify and delete, and environmental impact. A principles-based approach together with practical requirements to be released on a periodic basis that can keep up to date with emerging AI technologies, particularly given the rapid speed of advancement in this area, will provide clarity among organisations and agencies. This will enable the encouragement of innovation whilst protecting users and other stakeholders, creating certainty through guardrails and improving public trust in AI solutions. These principles should be able to be translated into effective assessment and assurance framework tools that organisations can embed into their risk assessment and ongoing monitoring processes.

The principles-based approach allows for flexibility, however, it will be important for the principles to be supported by examples in a non-exhaustive list which gives industry guidance on how the principles operate in practice. It is important to note that the EU and Canada have principals based that include lists of examples as part of their approach, and we should be aligning ourselves to best practice international regulatory models where possible. We do note that where legislation is highly prescriptive (for example Singapore) this can create challenges in operationalising the requirements.

Recommendation 4: *KPMG supports the adoption of a principles-based approach. In order to provide sufficient clarity and certainty, it will be important for the principles to be supported by examples in a non-exhaustive list which gives industry guidance on how the principles operate in practice.*

4. Are there high-risk use cases that government should consider banning in its regulatory response (for example, where there is an unacceptable level of risk)? If so, how should we define these?

KPMG Response Q4

KPMG considers that there are certain AI applications or Automated Decision Making (ADM) tools that may not be appropriate, that may be damaging and may undermine fundamental values of our society, including democracy and human rights, and should be banned.

While KPMG recommends against a technology-based regulatory approach in most cases, the EU AI Act is helpful in identifying what might be the criteria and implications of certain applications being considered high risk. Specifically, the EU AI Act prohibits certain AI systems because they present an unacceptable risk to human rights, public interests and human safety and dignity. Prohibited systems use subliminal or manipulative techniques to distort behaviour and cause harm, involve public social credit systems, expand facial recognition databases based on untargeted scraping of facial images, infer emotions in the workplace or education institutions, categorisation of people based on biometric data or biometric identification for law enforcement, except in limited circumstances in which such identification is permitted, and assessments of the likelihood an individual will commit a criminal offence based on certain traits. AI based surveillance applications or AI based assessments of a person's intent, social standing or character also need to be carefully considered given the high-risk of harm.

It is also important for Australia to be aligned with overseas approaches to banned applications so as not to become a haven for entities looking to utilise banned applications.

Recommendation 5: *KPMG considers that there are certain AI applications or Automated Decision Making (ADM) tools that may be damaging and may undermine fundamental values of our society, including democracy and human rights and, therefore, should be banned.*

5. Are the proposed principles flexible enough to capture new and emerging forms of high-risk AI, such as general-purpose AI (GPAI)?

KPMG Response Q5

KPMG considers that the principles are flexible enough to keep up with emerging technologies. At the same time, business commitment to strong AI ethics will help in controlling emerging AI technologies that may create significant harm. Australia's eight Artificial Intelligence Ethics Principles[2] which are designed to ensure AI is safe, secure and reliable, will be an important reference to ensure that the principles are flexible enough to keep up with emerging forms of AI.

In addition, while an AI application may not be high-risk, the use case of the AI application may create higher risks. Ensuring that entities undertake a risk assessment for each new use case will be important for controlling harm as new technology is created.

To ensure any regulatory framework keeps up with emerging technologies, we would also suggest a thorough government review is undertaken at least every two to three years and that any significant movement in international regulatory frameworks are considered and adopted as soon as practical.

Recommendation 6: *KPMG considers that the principles are flexible enough to keep up with emerging technologies. At the same time, business commitment to strong AI ethics will help in controlling for emerging AI technologies that cause significant harm. To ensure any mandatory guardrails or regulatory framework keeps up with emerging technologies, we suggest a thorough government led review is undertaken at least every two to three years and that any significant movement in international regulatory frameworks are considered and adopted as soon as practical.*

6. Should mandatory guardrails apply to all GPAI models?

KPMG Response Q6

KPMG considers that given the reach of general-purpose AI systems, mandatory guardrails should apply to these models. As per KPMG's previous submission on Safe and Responsible AI in Australia, it is important to ensure that unintended consequences and potential for harm are fully assessed and mitigated prior to, and during, the deployment of any general-purpose AI system. Particular care should also be given to human rights of vulnerable stakeholders (i.e., how they may use and/or be impacted by the outputs of these systems).

Additionally, Section 110 of the EU AI Act notes general-purpose AI models could pose systemic risks including actual or reasonably foreseeable negative effects in relation to major accidents, disruptions of critical sectors and serious consequences to public health and safety; actual or reasonably foreseeable negative effects on democratic processes, public and economic security; and the dissemination of illegal, false, or discriminatory content. The Act notes that systemic risks should be understood to increase with model capabilities and model reach, can arise along the entire lifecycle of the model, and are influenced by conditions of misuse, model reliability, model fairness and model security, the level of autonomy of the model, its access to tools, novel or combined modalities, release and distribution strategies, the potential to remove guardrails and other factors.

Recommendation 7: *KPMG considers that given the reach of general-purpose AI systems, mandatory guardrails should apply to these models. It may be useful to consider Section 110 of the EU-AI Act which outlines possible systemic risks of general-purpose AI models.*

7. What are suitable indicators for defining GPAI models as high-risk? For example, is it enough to define GPAI as high-risk against the principles, or should it be based on technical capability such

[2] Australia's AI Ethics Principles | Australia's Artificial Intelligence Ethics Framework | Department of Industry Science and Resources

as FLOPS (e.g. 10^25 or 10^26 threshold), advice from a scientific panel, government or other indicators?

KPMG Response Q7

Based on the EU AI Act, there are existing key indicators for defining high-risk GPAI models:

- **Risk-based regulation**: Higher-risk systems face stricter requirements, including conformity assessments and ex-ante conformity assessments.
- **High-impact capabilities**: A GPAI model is considered high-impact if it meets criteria such as being trained with vast compute power (more than (10^{25}) FLOPS) or having high-impact capabilities based on available indicators and benchmarks.
- **Systemic risk**: A model designated as having systemic risk for high-impact capabilities.
- Potential for Misuse: Advanced GPAI models are high-risk due to their potential for misuse, causing harm to people, community groups, and society at a wide scale and speed.
- **Compliance and governance**: High-risk GPAI models must comply with requirements like establishing a risk management system, performing data governance, developing technical documentation, automatic record-keeping, providing transparency, guaranteeing human oversight, and ensuring accuracy, robustness, and cybersecurity.

KPMG considers that any Australian framework should mirror the EU AI Act where practical.

Recommendation 8: *KPMG considers that any Australian framework should mirror the existing key indicators for defining high-risk GPAI models from the EU AI Act.*

Guardrails ensuring testing, transparency and accountability of AI

Consult questions

8. Do the proposed mandatory guardrails appropriately mitigate the risks of AI used in high-risk settings? Are there any guardrails that we should add or remove?

KPMG Response Q8

KPMG recommends that organisations need to consider mandatory guardrails when using AI in high-risk settings. There is an interlock between risk management processes and the handling of personal information and sensitive information. Both AI development and deployment present a variety of data issues, from a privacy and cybersecurity perspective (e.g. confidentiality, data usage rights, data quality, data sovereignty, data provenance, retention and destruction, etc.) particularly when AI is used in high-risk settings. For example, the use of poor-quality data can result in discrimination and bias within AI solutions which can result in significant adverse impacts on an individual, community or gender. There should be robust data governance processes and controls that are reviewed on an ongoing periodic basis, particularly where personal and sensitive information is used in connection with automated decision making. Additionally, there should be prescriptive guidance setting out the minimum requirements for high-risk AI solutions.

Practical requirements supporting mandatory guardrails for deploying high risk AI should be released particularly supporting guardrail one: *'Establish, implement and publish an accountability process including governance, internal capability and a strategy for regulatory compliance'* to promote active engagement from senior leadership to set the direction and articulate the values that will guide how the organisation uses AI.

KPMG supports guardrail 10 given that assurance mechanisms are shown to create trust in AI systems. Three out of four people (75 percent) report they would be more willing to trust AI systems when assurance mechanisms are in place that support ethical and responsible use.[3]

In relation to the operation of the assurance aspect of guardrail 10, there are varying approaches being adopted internationally. We note that the proposal paper considers that the conformity assessments could be carried out by the developers themselves, by a third-party or by government entities or regulators. The EU AI Act requires conformity assessments to be conducted by a notified body. A notified body is a conformity assessment body notified in accordance with the EU AI Act and other relevant EU harmonisation legislation, which performs third-party conformity assessment activities, including testing, certification, and inspection. According to the EU AI Act, notified bodies shall be independent of the provider of a high-risk AI system in relation to which they perform conformity assessment activities.[4] For guardrail 10 to be effective, it will be important for an effective assessment and assurance framework to be developed to ensure that these assessments are meaningful.

While we largely consider the guardrails in the proposal paper to be appropriate, the government should consider whether an additional guardrail is required that addresses system decommissioning. Currently, the guardrails do not address system decommissioning, so when a system comes to the end of its life, industry needs guidance as to how to most effectively ensure they can safely turn it off. Among the relevant considerations will be the safe removal of data.

While not directly related to the guardrails in the proposal paper, to reduce the potential harm of high-risk AI in government use cases, governments could look to mandate or preference suppliers that are accredited to ISO42001 when procuring AI. This could also be an interim measure before a formal conformity / assurance framework is developed.

Recommendation 9: *KPMG considers that the implementation of assurance mechanisms would facilitate greater trust in high-risk AI systems. For guardrail 10 to be effective, it will be important for an effective assessment and assurance framework to be developed to ensure that these assessments are meaningful.*

Recommendation 10: *KPMG recommends that the government consider whether an additional guardrail needs to be added to address the safe decommissioning of high-risk AI systems. Where high-risk AI*

[3] Trust in Artificial Intelligence | Global Insights 2023 - KPMG Australia
[4] Article 31: Requirements Relating to Notified Bodies | EU Artificial Intelligence Act

systems are not safely decommissioned, this could cause harm to any users of these systems including members of vulnerable populations.

Recommendation 11: *KPMG considers that there could be merit in governments mandating or preferencing suppliers that are accredited to a certain industry standard (for example ISO42001) when procuring AI. This could also be an interim measure before a formal conformity / assurance framework is developed for high-risk AI*

9. How can the guardrails incorporate First Nations knowledge and cultural protocols to ensure AI systems are culturally appropriate and preserve ICIP?
10. Do the proposed mandatory guardrails distribute responsibility across the AI supply chain and throughout the AI lifecycle appropriately? For example, are the requirements assigned to developers and deployers appropriate?
11. Are the proposed mandatory guardrails sufficient to address the risks of GPAI? How could we adapt the guardrails for different GPAI models, for example low-risk and high-risk GPAI models?

KPMG Response Q9-11

KPMG supports the applicability of all 10 guardrails across the AI supply chain and throughout the AI lifecycle. The guardrails support consistent practice in the adoption of AI in a safe and responsible way. This will give certainty to organisations and agencies about what developers and deployers of AI systems must do to comply with the guardrails. Practical requirements should be released to define and establish accountability across the AI supply chain.

12. Do you have suggestions for reducing the regulatory burden on small-to-medium sized businesses applying guardrails?

KPMG Response Q12

It will be important to ensure that adequate support is provided to businesses to help them prepare for regulatory changes, including education programs. KPMG suggests that bringing people together through appropriate professional bodies or forums such as the Council of Small Business Organisations Australia (COSBOA) or the Australian Small Business and Family Enterprise Ombudsman (ASBFEO) would be an efficient way to support businesses through the changes.

Another mechanism that could be considered is leveraging cloud-based platforms that could provide these protections as part of their services to small-to-medium sized businesses. These companies are often well placed to comply with regulatory obligations, stay up to date with advances in technology and regulation, and could use this service as a differentiator in the market.

Recommendation 12: *To reduce the regulatory burden on small-to-medium sized businesses applying guardrails, the Government could consider providing support to businesses to help them prepare for any regulatory changes, including education programs, and leveraging third-party AI providers who can provide protections as part of their services.*

Regulatory options to mandate guardrails

Consult questions

13. Which legislative option do you feel will best address the use of AI in high-risk settings? What opportunities should the government take into account in considering each approach?
14. Are there any additional limitations of options outlined in this section which the Australian Government should consider?
15. Which regulatory option/s will best ensure that guardrails for high-risk AI can adapt and respond to step-changes in technology?
16. Where do you see the greatest risks of gaps or inconsistencies with Australia's existing laws for the development and deployment of AI? Which regulatory option best addresses this, and why?

KPMG Response Q13-16

KPMG supports Option 3, a whole-of-economy approach which would involve introducing a new cross-economy AI-specific Act. KPMG has previously advocated for a framework-based approach, however, given the fast and significant advancements in technology, a specific AI Act may be a more appropriate regulatory model that would provide industry with the confidence it needs to invest and grow AI use cases and applications.

One of the key challenges for private and public organisations in the deployment of AI arises from the multiplicity of guidelines, frameworks, good practices and toolkits developed by the Australian Government as well as national and international policymakers. The development and adoption of simplified and interoperable legislation for AI should be accompanied by the identification of a leading regulatory body responsible for developing and enforcing AI legislation.

An AI-specific Act would provide clear and consistent expectations for those developing and deploying AI across the economy. Additionally, greater consistency with international regulatory frameworks would significantly reduce administrative burden, help with exporting technology out of Australia and set clearer expectations for the importation of technology. For example, Australian organisations that also operate in the EU, or have clients or data in the EU are already required to comply with the EU AI Act. Ensuring that Australia's approach is aligned with the EU AI Act would significantly reduce administrative burden for these organisations.

KPMG acknowledges the potential of added complexity and duplicate obligations with existing legislative frameworks with this option. Importantly, when developing the Act, it is critical to consider interoperability with existing frameworks such as privacy, discrimination and consumer laws, in order to minimise overlap and reduce regulatory burden. KPMG supports addressing duplication within the broader landscape of data-related regulatory requirements at the state and federal level. We encourage collaboration between Commonwealth agencies to ensure harmonisation between overlapping regulatory frameworks.

Option 2, a framework approach that adapts existing regulatory frameworks, could also be considered, which would provide flexibility, adaptability, and promote innovation. However, as outlined in the proposals paper, this option has limitations including retaining gaps across regimes and being limited to the scope and powers of current regulatory arrangements.

In relation to gaps in Australia's existing laws, many of Australia's current legislative frameworks that aim to address consumer and other individual harms are generally not yet adequately adapted to the use of AI and ADM technologies and their potential adverse impacts. For example, we note that the current Copyright Act 1968 does not include or consider AI in its scope.

Recommendation 13: *KPMG supports introducing a new cross-economy AI-specific Act that provides clear and consistent expectations on those developing and deploying AI, better enables interoperability with international approaches, and addresses complexity and duplication within existing legislative frameworks..*

Recommendation 14: *Greater consistency with international regulatory frameworks would significantly reduce administrative burden, help with exporting technology out of Australia, and set clearer expectations for the importation of technology.*

Key authors and contacts

Kelly Henney
National Leader, Privacy & Data Protection

James Mabbott
Partner in Charge, KPMG Futures

Merriden Varrall
Partner, Geopolitics

Jon Berry
Associate Director, Geopolitics

Francine Hoo
Director, Enterprise Advisory

Shubham Singhal
Director, Privacy & Data Protection

Leah Mooney
Director, Governance, Risk & Compliance

Niran Garcha
Privacy & Data Protection, Manager

Sarah Minahan
Director, ESG Assurance & Advisory

Jessica Wyndham
Trusted AI Lead, KPMG

Samantha Hamilton
Manager, Governance, Risk & Compliance

Judith Mendes
Senior Consultant, ESG Assurance & Advisory