Avant
By doctors for doctors

Avant Mutual Group Limited
ABN 58 123 154 898

Registered office
Level 6, Darling Park 3
201 Sussex Street, Sydney NSW 2000

Postal address
PO Box 746 Queen Victoria Building
Sydney NSW 1230

DX 11583 Sydney Downtown

avant.org.au

Telephone 02 9260 9000 **Fax** 02 9261 2921
Freecall 1800 128 268 **Freefax** 1800 228 268

4 October 2024

Via the Consultation Hub

Introducing mandatory guardrails for AI in high risk settings

Thank you for the opportunity to provide a response to the Department of Industry, Science and Resources ("**DISR**") consultation on Safe and Responsible Artificial Intelligence ("**AI**") in Australia – proposal for introducing mandatory guardrails for AI in high risk settings.

Our submission is attached.

Please contact me or Tracy Pickett on the details below if you require any further information or clarification of the matters raised in the submission.

Yours sincerely

Georgie Haysom
General Manager, Advocacy, Education and Research
Email: georgie.haysom@avant.org.au

Tracy Pickett
Legal & Policy Advisor
Email: tracy.pickett@avant.org.au

Avant Mutual Group Limited
ABN 58 123 154 898

Avant submission to the consultation on introducing mandatory guardrails for AI in high risk settings

Avant is a member-owned doctors' organisation and Australia's largest medical indemnity insurer, committed to supporting a sustainable health system that provides quality care to the Australian community. Avant provides professional indemnity insurance and legal advice and assistance to about 90,000 healthcare practitioners and students around Australia (more than half of Australia's doctors). Our members are from all medical specialities and career stages and from every state and territory in Australia.

We assist members in civil litigation, professional conduct matters, coronial matters and a range of other matters. Our Medico-legal Advisory Service provides support and advice to members and insured medical practices when they encounter medico-legal issues. We aim to promote quality, safety and professionalism in medical practice through advocacy, research and medico-legal education.

General comments

Avant welcomes increased governance in relation to the implementation of safe and responsible AI in Australia and acknowledges the delicate balance required between regulation for safety and consistency, and promotion of innovation.

While we acknowledge that this consultation adopts a whole of economy approach, our focus is on the need to promote safe and clinically appropriate regulation of AI in the context of health care. We will be making submissions to sector-specific consultations conducted by the Department of Health and Aged Care and the Therapeutic Goods Administration.

Avant supports the need for clearer frameworks and guidelines to address the complexities of AI in healthcare, ensuring that responsibility and liability are properly managed and that both doctors and patients are adequately protected.

In summary, our key points are:

- We are broadly supportive of the introduction of guardrails as outlined by DISR and we agree that they should be mandatory in high-risk settings.
- We agree that the use of AI in health care is high-risk, based on the potential for adverse impacts on an individual's physical or mental health or safety. However, there is a scale of risk for the AI products that might be used in the health care setting.
- We acknowledge that existing regulatory frameworks are ill-equipped to address the unique challenges posed by using AI in health care. Current regulations may not sufficiently cover the safety, efficacy, and ethical use of AI, and will need to be adapted.
- The regulatory settings for AI products used in health care should be commensurate with the scale of risk associated with the type of product. Not all uses need to be

treated in the same way. The regulation of some products such as AI scribes may also need to be reconsidered.

- Clinicians may be both deployers and users of AI in settings such as medical practices. It is essential that there is clear guidance about the obligations on clinicians and how to comply with the guardrails to limit regulatory burden and enable adoption.
- We agree with the preventive approach to accountability and risk management, but if harm does occur, there should be mechanisms for appropriately determining liability and obtaining redress, beyond the complaints handling processes outlined in relation to guardrail 7.
- Liability should be appropriately apportioned and should not be shifted away from those who are able to control the risks, for example, through contractual indemnity clauses and disclaimers.
- The guardrails must also allow for appropriate insurance and indemnity arrangements for developers, deployers and users in the health care sector so that there is accountability and responsibility for patient safety across the AI lifecycle. Both practitioners and patients alike need to have the confidence that insurance issues can be appropriately addressed should they arise.
- Further consideration is required about the best regulatory approach to implement the mandatory guardrails so that they are applied consistently and without duplication. We broadly support having a principles-based framework, supported by guidance on how to apply the principles in practice which might be sector specific. A relevant comparison is Australia's privacy regime model.
- Ongoing consultation will be essential and we recommend that sector-specific working groups are formed, made up of key stakeholders including insurers.

Responses to consultation questions

Defining High Risk AI

1. Do the proposed principles adequately capture high risk AI?

Yes.

Are there any principles we should add or remove?

No.

Please identify any:
- low risk cases that are unintentionally captured
- categories of uses that should be treated separately such as for defence or national security purposes

No response.

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

No response.

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?

Yes – the principles give enough clarity and certainty.

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

We would welcome the inclusion of examples of how to apply the principles to use cases involving various settings and AI. The examples provided in the proposals paper are effective as they explore borderline settings and explain both what is and isn't considered high-risk in these settings.

How can this list capture emerging uses of AI?

No response.

4. Are there high risk cases that government should consider banning in its regulatory response (for example, where there is an unacceptable level of risk)?

Yes.

If yes – how should we define these?

High risk cases should be defined with regard to Australia's AI Ethics Principles. This would ensure interoperability with international principles such as those identified in the EU AI Act.

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

Yes.

Please provide any additional comments.

No response.

6. Should mandatory guardrails apply to all GPAI models?

Yes.

Please provide any additional comments.

Mandatory guardrails should apply to all GPAI models because there is the potential for these models to be used for a wide range of purposes and the possible applications and risks cannot be foreseen.

7. What are suitable indicators for defining GPAI models as high risk?

No response.

Guardrails ensuring testing, transparency and accountability of AI

8. Do the proposed mandatory guardrails appropriately mitigate the risks of AI used in high risk settings?

Yes.

Please provide any additional comments.

We agree that the introduction of the proposed mandatory guardrails is an appropriate means of responding to the risks of AI in high-risk settings. The guardrails themselves are comprehensive, addressing the AI lifecycle and the AI supply chain. The principles-based approach is also effective, enabling consistent application across the whole economy as well as sector-specific application and responses.

We also agree that the use of AI in health care is high-risk, based on the potential for adverse impacts on an individual's physical or mental health or safety. These additional sensitivities should be considered so that AI is adopted safely and in a way that enhances public and professional trust.

However, there is a scale of risk for the AI products that might be used in the health care setting. Uses in healthcare range from providing and collecting information, triage and prioritisation, document summarisation, prediction and clinical decision support, to diagnosis, treatment and ongoing clinical management. Our view is that the regulatory settings for AI products used in health care should be commensurate with the scale of risk associated with the type of product. Not all uses need to be treated in the same way. The regulation of some products which are currently unregulated, such as AI scribes for clinical notetaking, may also need to be reconsidered. We will provide further commentary in our submission to the Department of Health and Aged Care consultation.

Given that the mandatory guardrails are to apply to developers and deployers of high-risk AI systems across the AI lifecycle, we support the following requirements to mitigate the risks:

- The ongoing obligation on developers to monitor and refine their AI system once deployed (guardrail 4)
- Deployers to inform developers of any adverse incidents or risks that emerge while the system is in use (guardrail 8)
- The requirement on developers to undertake a new conformity assessment if the AI system changes in a way that affects compliance with the guardrails (guardrail 10). This should not be the responsibility of deployers as implied in the table on page 67 of the proposals paper.

While the proposals paper indicates that the guardrails will operate "*alongside laws and regulatory instruments to hold organisations accountable when harm does occur*", we look forward to considering the details of how this will operate in practical terms.

Ongoing consultation will be essential and we recommend that sector-specific working groups are formed, made up of key stakeholders including insurers.

Are there any guardrails that we should add or remove?

No.

9. How can the guardrails incorporate First Nations knowledge and cultural protocols to ensure AI systems are culturally appropriate and preserve Indigenous Cultural and Intellectual Property?

No response.

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?

Yes.

Please provide any additional comments.

We make two comments from a health care perspective in relation to responsibility and assigned requirements in the proposed guidelines:
1) responsibility, liability and insurance considerations
2) responsibility for conformity assessments.

Responsibility, liability and indemnity considerations
Avant supports a risk-based approach to promote the apportionment of liability and responsibility to those who have the ability to affect outcomes and control the risk. In particular, we support the intention that the mandatory guardrails require developers and deployers of high-risk AI to take specific steps across the AI lifecycle to:

- test to ensure that the systems perform as intended and meet appropriate performance metrics
- ensure transparency in relation to product development and use
- promote accountability for governing and managing the risks of AI systems.

Currently, there is legal uncertainty around accountability, responsibility and liability across the AI lifecycle in healthcare. It can be difficult to determine who is accountable when AI-driven decisions result in errors or adverse outcomes. The complexity of AI algorithms

makes it challenging to trace the decision-making process, creating medico-legal risks and complicating the implementation of risk management processes and the assignment of liability.

There is a risk that responsibility for the design or function of AI could be unfairly shifted onto doctors. This has been exacerbated by the presence of broad indemnity clauses in some AI provider contracts, which attempt to absolve these companies of responsibility and liability and place the burden on doctors using the AI system. This not only gives rise to moral hazard, but also it exposes doctors to legal risk and raises concerns about whether their professional indemnity insurance will cover liability under such contracts.

The guardrails must also allow for appropriate insurance and indemnity arrangements for developers, deployers and users in the health care sector so that there is accountability and responsibility for patient safety across the AI lifecycle. Both practitioners and patients alike need to have the confidence that insurance issues can be appropriately addressed should they arise.

The challenge of ensuring proper accountability is amplified by the need for accessible and historical data on AI performance. If the data and operational information the AI systems rely upon are not properly recorded, stored and maintained (as required under guardrail 9), it becomes extremely difficult to trace the source of the error that has caused patient harm and determine liability at a later date. Guardrail 9 should encompass the need to keep historical data.

In guardrail 7, the discussion refers to complaints handling. However, redress for harm is broader than that and includes appropriate compensation. We note also that the table on the last page the box outlining what this means for developers is blank.

We recommend the following:

- That developers of high-risk AI software are prevented from contracting out of their responsibilities under the guardrails. This may prevent those who have the capacity to reduce risk from avoiding responsibility through indemnity clauses.
- Developers and/or deployers not based within Australia should be required to have an Australian base or nexus to ensure that they will fulfil their accountability and reporting requirements under the guardrails.
- That standards be adopted to ensure that the guardrail obligations for developers and deployers remain in force for the period that the developer or deployer maintains their involvement with the AI system. Developers and deployers should be required to ensure they have insurance to cover future liability. This is particularly important in the health sectors given limitation periods under civil liability legislation.
- In guardrail 4, standards should be adopted to identify relevant time frames for responding to feedback from deployers about changes to the AI system. Failing to provide some guidance on the time frames may result in delays and increased risk.
- Guardrail 9 should encompass the need to keep historical data for relevant time periods, given limitation periods under civil liability legislation.

Responsibility for conformity assessments
We agree that developers must undertake a conformity assessment before their AI system is deployed, as outlined in guardrail 10. This should also be required of developers when the AI system is retrained or changes in a way that affects compliance with the guardrails. This should not be the responsibility of deployers as implied in the table on page 67 of the proposals paper.

Responsibility for performing the conformity assessment is currently unclear. In our view conformity assessments should be performed by third-parties or government entities as this would be more independent and objective than self-assessment. This would need to be supported by the requirement to maintain records and a documented accountability process as outlined in the proposals paper.

11. Are the proposed mandatory guardrails sufficient to address the risks of GPAI?

Yes.

If yes, please provide any additional comments.

Avant supports mandatory guardrails for all GPAI models to protect against unforeseeable risks to users and the public at large.

12. Do you have any suggestions for reducing regulatory burden on small-to-medium sized businesses applying guardrails? Please provide any additional comments.

Yes.

Clinicians may be both deployers and users of AI in settings such as medical practices, which are often small businesses. The guardrails will create many obligations on clinicians, and with it a significant regulatory burden. The accountability processes suggested under guardrail 1 are an example of this. There will be a barrier to adoption if the regulatory and compliance burden is too great.

We recommend the following:
- It is essential that there is clear guidance about the obligations on clinicians and how to comply with the guardrails to limit any regulatory burden and enable adoption.
- Guidance should acknowledge the potential overlap between a clinician's obligations as a deployer and a user of AI in health care.
- Under guardrail 9, regulatory and compliance requirements should align with existing practice standards for clinicians (e.g. RACGP accreditation standards).
- Requirements should also integrate into existing risk management processes wherever possible.

Regulatory options to mandate guardrails

13. Which legislative option do you feel will best address the use of AI in high-risk settings?

No response.

What opportunities should the government take into account in considering each approach?

See response to Q15 below.

14. Are there any additional limitations of options outlined in this section which the Australian government should consider?

See response to Q15 below.

15. Which regulatory option(s) will best ensure that guardrails for high-risk AI can adapt and respond to step-changes in technology?

Other.

Further consideration is required about the best regulatory approach to implement the mandatory guardrails so that they are applied consistently and without duplication across and within sectors.

We broadly support having a principles-based framework, accompanied by guidance on how to apply the principles/guardrails in practice which might be sector specific. A relevant comparison is the Commonwealth privacy regime model.

We also believe that the regulatory regime should not be overly burdensome nor be a barrier to innovation or adoption. However, the balance between regulation and innovation needs to be appropriate to achieve this.

Option 1 – the domain specific approach – could be beneficial because:

- It leverages the existing economy wide and sector specific laws to apply the guardrails.
- Adoption of the guardrails in the context of existing legislation capitalises on the associated familiarity and knowledge of the legislation.
- It will allow for the introduction of the guardrails at the earliest possible opportunity, supported by tailored regulation for each specific area.

However, we agree that option 1 is likely to encourage gaps between regulated sectors and regulatory silos.

We agree that option 2 – the framework approach – would be more flexible and able to respond to step-changes in technology but also has limitations and, like option 1, could lead to regulatory gaps and silos.

16. Where do you see the greatest risks of gaps and inconsistencies within Australia's existing laws for the development and deployment of AI?

- There are scales of "high-risk" depending on the nature of the risk and the sector in which it arises. How these are treated in regulation across sectors may create inconsistencies.
- If options 1 or 2 are adopted, there is a risk of inconsistency between Commonwealth and state and territory legislation. Depending on the sector within which the AI operates and the location, there may be differing regulatory responses.
- There may be gaps in applying privacy legislation to specific AI use cases and we note that the privacy legislation is under review in this regard.
- There may be gaps in product liability laws, specifically whether software systems meet the definition of a product. This should be considered as part of the priority consumer law review.
- There may be a gap in ensuring appropriate responses (such as override or shut down) if an AI system is found to be an unacceptable risk once deployed.
- There is a need to ensure that there are appropriate and harmonised penalty provisions for non-compliance.

Avant Mutual
4 October 2024