Advanced AI Demands More than the Standard Regulatory Playbook

Response to the Productivity Commission interim report on "Harnessing data and digital technology"

15 September 2025

About Gradient Institute

We are an independent, nonprofit Australian research institute and registered charity that works to build safety, ethics, accountability and transparency into AI systems. We conduct technical research, education and technical assessments, as well as offer technical guidance for AI policy development. Find more about us at https://www.gradientinstitute.org/ and contact us at info@gradientinstitute.org.

Executive Summary

This submission contends that the Productivity Commission's (PC) draft recommendations on Artificial Intelligence (AI) are fundamentally flawed because they apply a generic regulatory framework that is agnostic to technology capability and that implicitly treats *all* AI as a conventional technology. Whilst principles such as conducting "gap analyses" and prioritising "technological neutrality" are standard regulatory best practice and could be appropriate for some AI technologies, their application here ignores the unique and exceptional nature of *advanced* AI, making the framework unsuitable and potentially unsafe.

Drawing on scientific consensus from the International AI Safety Report 2025, we argue that the nature of *advanced* AI is profoundly different from past technologies (including some AI technologies) in a highly consequential manner. It is characterised by a unique combination of three properties: (1) its core capabilities are evolving at an unpredictable pace, (2) its general-purpose nature means risks cannot be managed solely at the use-case level and can emerge in unanticipated ways, and (3) there is a fundamental lack of a scientific understanding about how these systems work. These properties directly challenge the core assumptions underpinning the PC's incremental approach, which presupposes a level of predictability and comprehensibility to all forms of AI that the evidence shows does not exist for advanced AI.

Applying a standard regulatory playbook to what science indicates could be a paradigm-shifting technology is not a prudent course of action. Therefore, we recommend a shift in strategy from a presumption of continuity to one based on navigating deep uncertainty. The most urgent and constructive path forward is not necessarily to impose broad, mandatory regulations today, but to make a foundational national investment in the sovereign scientific and technical capability required to inform sensible government policy at this critical moment. We urge the government to establish a national body tasked with researching, monitoring, and measuring AI's evolving capabilities and risks. This will create the necessary awareness, knowledge and capability to appropriately inform government policy to harness the immense productivity benefits of AI and manage its profound risks with foresight and wisdom.

Scope of this submission

This submission provides feedback on the Productivity Commission's (PC) interim report, "Harnessing data and digital technology," published in August 2025.[1] We thank the PC for its work and welcome the opportunity to comment on it. Our feedback focuses specifically on Chapter 1, which addresses Artificial Intelligence (AI) and its associated Draft Recommendations 1.1, 1.2, and 1.3.

The PC's recommendations on AI are not about AI

The core principles embedded in the PC's draft recommendations on AI[2] – that regulation should be "proportionate, risk-based, outcomes-based and technology-neutral where possible," that "gap analyses" should precede new rules, and that technology-specific regulation must be a "last resort" – are consistent with modern regulatory best practice in Australia.[3] They are principles of a "standard regulatory playbook".

To illustrate their universality, consider a substitution exercise. If the term "AI" were replaced with a technology that, although new, had fixed capabilities, was well understood scientifically, and whose impact was controllable through its context of use, how would the guiding principles of the recommendations be any different? They would not: if we were talking about a regulatory system for a new form of transport (for example) one would want a proportionate, risk-based, and outcomes-based approach; one would conduct a gap analysis before implementing new rules, and reserve technology-specific regulation as a last resort.

This is precisely the issue. These principles work for more conventional technology (including some forms of AI technologies). This means they are consistent with a premise that *all* of AI technology is "sufficiently conventional," at least for regulatory purposes. This is manifested in the following assumptions implied by the recommendations (which make no distinctions between different forms of AI and therefore presumably are intended to apply to all):

---

[1] https://www.pc.gov.au/inquiries/current/data-digital#interim
[2] The PC draft recommendations on AI are reproduced in the Appendix of this submission.
[3] Australian Government Guide to Policy Impact Analysis.
https://oia.pmc.gov.au/resources/guidance-impact-analysis/australian-government-guide-policy-impact-analysis

- That the dynamics of the evolution, adoption, and impacts of all forms of AI will be sufficiently incremental to allow for a "gap analysis" process to unfold at a pace consistent with that of existing regulatory institutions.
- That existing policymakers, regulators, and courts possess the scientific and technical capacity to comprehend the capabilities of all forms of AI to properly assess risk.
- That the societal impact of all forms of AI can be adequately addressed primarily through a use-case-specific framework.

Crucially, the recommendation to pause mandatory guardrails presupposes that this incremental evidence-gathering posture is safe, or that the interim risks of delaying proactive regulation are acceptable. The framework does not account for the possibility that advanced forms of AI constitute a potential paradigm shift sufficient to justify questioning established principles of governance, rather than a manageable innovation that fits within them.

The scientific case for Advanced AI's exceptionalism

The critical question we ask is not whether the PC's principles are sound in general, but whether they remain sufficient for *advanced AI* in particular. The starting point is to note that the interim report does not differentiate between different types of AI, and yet there are fundamental and consequential differences between more traditional types of AI (such as machine learning models in use for credit scoring and marketing for decades) and the types of AI that the [International AI Safety Report 2025](https://www.gov.uk/government/publications/international-ai-safety-report-2025) (IAISR'25) calls *Advanced AI*: "the most advanced general-purpose AI systems at the time of writing, as well as future systems that might be even more capable".[4] Evidence from the IAISR'25 suggests that advanced AI exhibits a unique combination of characteristics that profoundly challenge the suitability of the standard regulatory playbook that the PC has recommended for all forms of AI.[5] The focus and scope of our submission is specifically on Advanced AI.

High unpredictability in the pace and nature of capability change. The PC's "gap analysis" approach implicitly assumes a pace of AI advancement and adoption that can be managed within standard regulatory timelines. The IAISR'25 directly challenges this in the case of advanced AI, stating that future capability advancements could be "anything from slow to extremely rapid." This is not about the speed of technology adoption, but the speed with which the technology itself morphs into a more capable version. This contradicts the implied

---

[4] International AI Safety Report 2025 (IAISR'25)
https://www.gov.uk/government/publications/international-ai-safety-report-2025
[5] The interim report refers to IAISR'25 but does not engage with the evidence that challenges the AI recommendations made.

assumption in the interim report that the evolution of all forms of AI will be slow enough for standard regulatory processes to safely keep pace. While that might be true, the scientific evidence shows we simply don't know what's coming – advancements could be "extremely rapid". The IAISR'25 provides substantial evidence for this possibility, drawing both from the empirical trajectory observed over recent years – particularly in the last five years with rapid developments in large language models powering technologies like ChatGPT – and from the scientific understanding of what drives that trajectory.

General-purpose nature. The PC's approach appears to assume that risk management for AI can be conducted primarily at the "specific use-case" level. For narrow AI systems, this can be a sensible assumption. However, it no longer holds in general for the general-purpose AI systems driving the modern AI revolution. The IAISR'25 notes that the "breadth of use cases complicates safety assurance" because these systems are "being used for many (often unanticipated) tasks in many contexts." Furthermore, how to accommodate new use cases that emerge after the gap analysis? This challenges the assumption that a use-case-centric framework is adequate for advanced AI.

Fundamental scientific uncertainty. The reliance on existing legal and regulatory structures to assess risk assumes they have the scientific and technical capability to do so. The IAISR'25 makes it clear, however, that core AI capabilities are not well understood scientifically. It states that even developers "understand little about how their models operate internally." This lack of firm scientific understanding poses a fundamental challenge to risk assessment. If developers cannot fully explain a model's behaviour, it is difficult to see how external regulators can be expected to reliably assess its risks, refuting the assumption that existing institutions have the required technical capacity.

How many technologies in history have simultaneously evolved at an unpredictable pace, defied containment to specific use cases, and remained fundamentally unintelligible even to their creators?

Rethinking how to regulate in light of AI

The scientific evidence strongly suggests that advanced AI warrants exceptional treatment. The confluence of uncertainty about both the speed and nature of capability increases, an inherently general-purpose nature, and deep scientific uncertainty about how it works makes

advanced AI qualitatively different from prior technologies in highly consequential ways. A regulatory framework for AI in Australia must grapple with these realities, not ignore them.

By applying a generic regulatory framework that is agnostic to technology capability, the PC's report implicitly assumes a level of incrementalism, specificity, and comprehensibility that the scientific evidence indicates does not exist for advanced AI. This suggests the very foundation of how to approach regulation must be reconsidered. In order to address the risks posed by advanced AI, there is a need to move from a framework designed for typical technological uncertainty to one built for navigating deep uncertainty.

How can Australia safely harness the benefits of advanced AI?

A modern, agile, and evidence-based approach to advanced AI governance must be built on a foundation of scientific understanding. We advise the government to:

1. **Acknowledge the scientific evidence of advanced AI's unprecedented exceptionalism**. The starting point for any robust national strategy for AI policy must be the scientific reality. The [International AI Safety Report 2025](#) provides a clear, consensus-based foundation for this view with regards to advanced AI, highlighting properties – high uncertainty in capability change, general-purpose applicability, and fundamentally limited scientific understanding – that directly challenge the standard ways of regulating in face of typical technological change.

2. **Invest immediately and substantially in sovereign scientific and technical state capacity**. The primary policy imperative is not to regulate prematurely, but to build the capability to support scientifically grounded government policy and ensure regulatory powers to act or intervene quickly where necessary. Primarily, Australia must invest in a national body with a clear mission: to conduct foundational scientific and technical research into AI, its capabilities and safety, and its societal impacts. The goal is to reduce uncertainty, better understand its nature, and create the awareness, knowledge and capability that enables appropriate evidence-based policy – both preemptive and responsive. This should include not only regulatory policy, but also industrial policy and beyond: investing in workforce upskilling and high speed internet won't be enough.[6] This investment, and making Australia a real protagonist in the global effort to understand this technology and its implications, is the most direct step

---

[6] On page 10 of its report, under the heading "Governments have a key role in making the most of the AI Opportunity," the PC outlines its recommended government investments for AI in Australia. Beyond its proposed regulatory approach, the report essentially only addresses workforce upskilling and high-speed internet access.

to ensuring AI development and use is safe and aligned with Australia's national interest and economic productivity.

3. **Task this new capacity to inform adaptive governance and government policy more generally**. This sovereign capability should have a clear remit to assist with regulatory action and, where appropriate, reform designed for an era of rapid technological change. It should support the development of frameworks that are more adaptive, anticipatory, and robust to deep uncertainty than standard regulatory principles would suggest. This will allow Australia to move beyond a reactive "gap analysis" posture toward a more forward-looking, strategic approach to both AI and its governance.

Conclusion

The Productivity Commission has offered a set of prudent, well-established principles for regulating in light of conventional technological change – but has applied it to *all forms of AI*, including the kind of advanced AI that is rapidly developing, general purpose, which isn't well understood scientifically, and which could represent a paradigm shift. While standard principles are sound for conventional technologies, they are based on assumptions of sufficient predictability and comprehensibility that are fundamentally challenged by the scientific consensus on advanced AI. Applying a standard regulatory playbook to what could plausibly be the most transformative and poorly understood technological transition in history is not a safe or productive path.

For Advanced AI, the most constructive and urgent way forward is not necessarily to impose broad, mandatory regulations today. Rather, it is to make a foundational national investment in the scientific and technical capacity required to understand this technology and its impacts. By building a sovereign scientific capability to rigorously assess and continuously monitor Advanced AI, its capabilities and risks, Australia can create the foundation to appropriately inform a truly adaptive and effective regulatory regime – one that can both seize the immense productivity benefits of AI and manage its profound risks with foresight and wisdom.

Appendix: The Productivity Commission's draft recommendations on AI

Draft recommendation 1.1.

Productivity growth from AI will be built on existing legal foundations. Gap analyses of current rules need to be expanded and completed.

Australian governments play a key role in promoting investment in digital technology, including AI, by providing a stable regulatory environment. Any regulatory responses to potential harms from using AI must be proportionate, risk-based, outcomes-based and technology-neutral where possible. The Australian Government should continue, complete, publish and act on ongoing reviews into the potential gaps in the regulatory framework posed by AI as soon as possible.Where relevant gap analyses have not begun, they should begin immediately. All reviews of the regulatory gaps posed by AI should consider:

• the uses of AI

• the additional risk of harm posed by AI (compared to the status quo) in a specific use-case

• whether existing regulatory frameworks cover these risks potentially with improved guidance and enforcement; and if not how to modify existing regulatory frameworks to mitigate the additional risks.

Draft recommendation 1.2

AI-specific regulation should be a last resort

AI-specific regulations should only be considered as a last resort for the use cases of AI that meet two criteria. These are:

• where existing regulatory frameworks cannot be sufficiently adapted to handle the issue

• where technology-neutral regulations are not feasible.

Draft recommendation 1.3

Pause steps to implement mandatory guardrails for high-risk AI

The Australian Government should only apply the proposed 'mandatory guardrails for high-risk AI' in circumstances that lead to harms that cannot be mitigated by existing regulatory frameworks and where new technology-neutral regulation is not possible. Until the reviews of the gaps posed by AI to existing regulatory structures are completed, steps to mandate the guardrails should be paused.