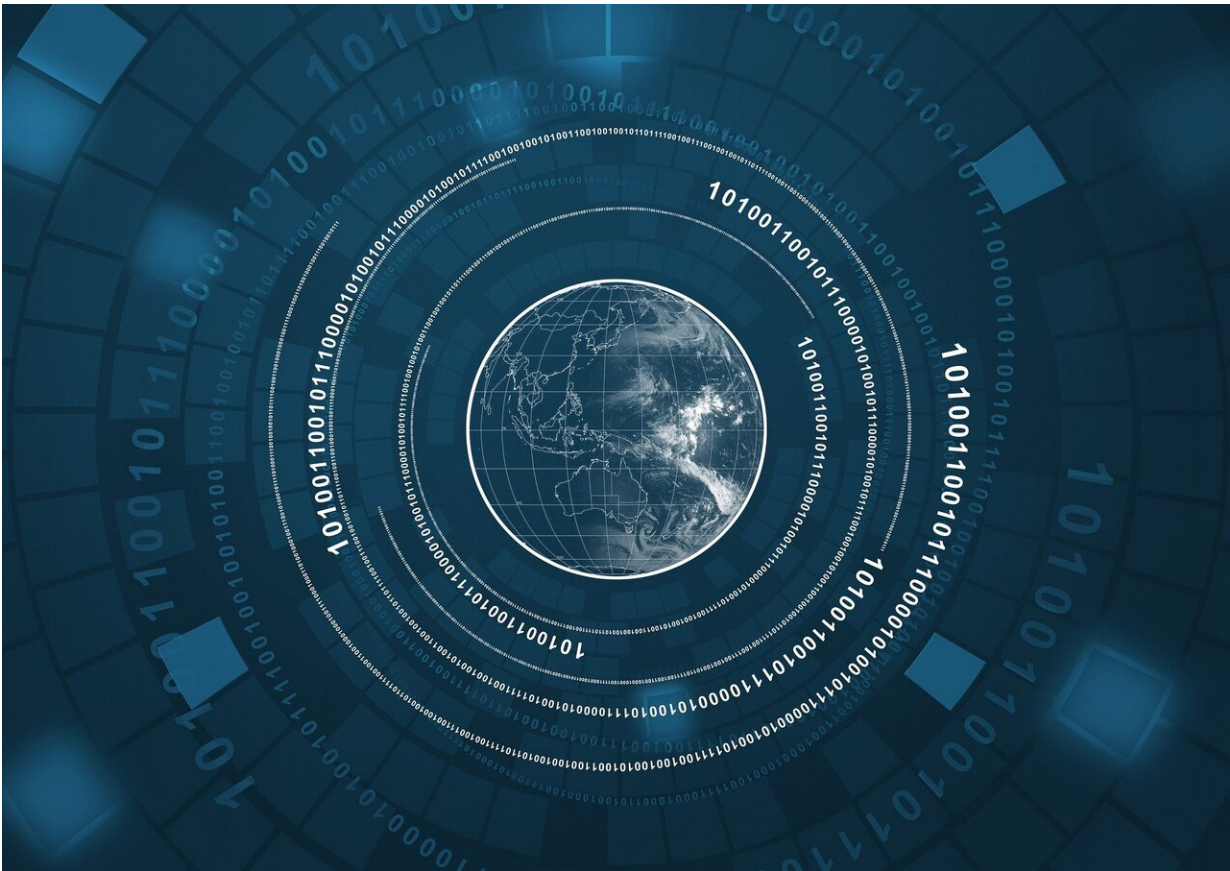# How Australia's new AI 'guardrails' can clean up the messy market for artificial intelligence

September 5 2024, by Nicholas Davis

Credit: CC0 Public Domain

Australia's federal government has today launched a proposed set of

[mandatory guardrails for high-risk AI](#) alongside a [voluntary safety standard](#) for organizations using AI.

Each of these documents offers ten mutually reinforcing guardrails that set clear expectations for organizations across the AI supply chain. They are relevant for all organizations using AI, including internal systems aimed at boosting employee efficiency and externally-facing systems such as chatbots.

Most of the guardrails relate to things like accountability, transparency, record-keeping and making sure humans are overseeing AI systems in a meaningful way. They are aligned with emerging international standards such as the [ISO standard for AI management](#) and the [European Union's AI Act](#).

The proposals for mandatory requirements for high-risk AI—which are open to [public submissions](#) for the next month—recognize that AI systems are special in ways that limit the ability of existing laws to effectively prevent or mitigate a wide range of harms to Australians. While defining precisely what constitutes a high-risk setting is a core part of the consultation, the proposed principle-based approach would likely capture any systems that have a legal effect. Examples might include AI recruitment systems, systems that may limit human rights (including some facial recognition systems), and any systems that can cause physical harm, such as autonomous vehicles.

Well-designed guardrails will improve technology and make us all better off. On this front, the government should accelerate law reform efforts to clarify existing rules and improve both transparency and accountability in the market. At the same time, we don't need to—nor should we—wait for the government to act.

The AI market is a mess

As it stands, the market for AI products and services is a mess. The central problem is that people don't know how AI systems work, when they're using them, and whether the output helps or hurts them.

Take, for example, a company that recently asked for my advice on a generative AI service projected to cost hundreds of thousands of dollars each year. It was worried about falling behind competitors and having difficulty choosing between vendors.

Yet, in the first 15 minutes of discussion, the company revealed it had no reliable information around the potential benefit for the business, and no knowledge of existing generative AI use by its teams.

It's important that we get this right. If you believe even a fraction of the hype, AI represents a huge opportunity for Australia. [Estimates referenced by the federal government](#) suggest the economic boost from AI and automation could be up to A\$600 billion every year by 2030. This would lift our GDP to 25% above 2023 levels.

But all of this is at risk. The evidence is in the alarmingly high failure rates of AI projects ([above 80% by some estimates](#)), an array of reckless rollouts, [low levels of citizen trust](#) and the prospect of thousands of Robodebt-esque crises across both industry and government.

The information asymmetry problem

A lack of skills and experience among decisionmakers is undoubtedly part of the problem. But the rapid pace of innovation in AI is supercharging another challenge: information asymmetry.

Information asymmetry is a simple, [Nobel prize-winning economic concept](#) with serious implications for everyone. And it's a particularly

pernicious challenge when it comes to AI.

When buyers and sellers have uneven knowledge about a product or service, it doesn't just mean one party gains at the other's expense. It can lead to poor-quality goods dominating the market, and even the market failing entirely.

AI creates information asymmetries in spades. AI models are technical and complex, they are often embedded and hidden inside other systems, and they are increasingly being used to make important choices.

Balancing out these asymmetries should deeply concern all of us. Boards, executives and shareholders want AI investments to pay off. Consumers want systems that work in their interests. And we all want to enjoy the benefits of economic expansion while avoiding the very real harms AI systems can inflict if they fail, or if they are used maliciously or deployed inappropriately.

In the short term, at least, companies selling AI gain a real benefit from restricting information so they can do deals with naïve counterparties. Solving this problem will require more than upskilling. It means using a range of tools and incentives to gather and share accurate, timely and important information about AI systems.

What businesses can do today

Now is the time to act. Businesses across Australia can pick up the [Voluntary AI Safety Standard](#) (or the [International Organization for Standardization's version](#)) and start gathering and documenting the information they need to make better decisions about AI today.

This will help in two ways. First, it will help businesses to take a structured approach to understanding and governing their own use of AI

systems, to ask useful questions to (and demand answers from) their technology partners, and to signal to the market that their AI use is trustworthy.

Second, as more and more businesses adopt the standard, Australian and international vendors and deployers will feel market pressure to ensure their products and services are fit for purpose. In turn, it will become cheaper and easier for all of us to know whether the AI system we're buying, relying on or being judged by actually serves our needs.

Clearing a path

Australian consumers and businesses both want AI to be safe and responsible. But we urgently need to close the huge gap that exists between aspiration and practice.

The National AI Centre's [Responsible AI index](#) shows that while 78% of organizations believed they were developing and deploying AI systems responsibly, only 29% of organizations were applying actual practices towards this end.

Safe and responsible AI is where good governance meets good [business](#) practice and human-centered technology. In the bigger picture, it's also about ensuring that innovation thrives in a well-functioning market. On both these fronts, standards can help us clear a path through the clutter.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

intelligence (2024, September 5) retrieved 6 September 2024 from
https://techxplore.com/news/2024-09-australia-ai-guardrails-messy-artificial.html