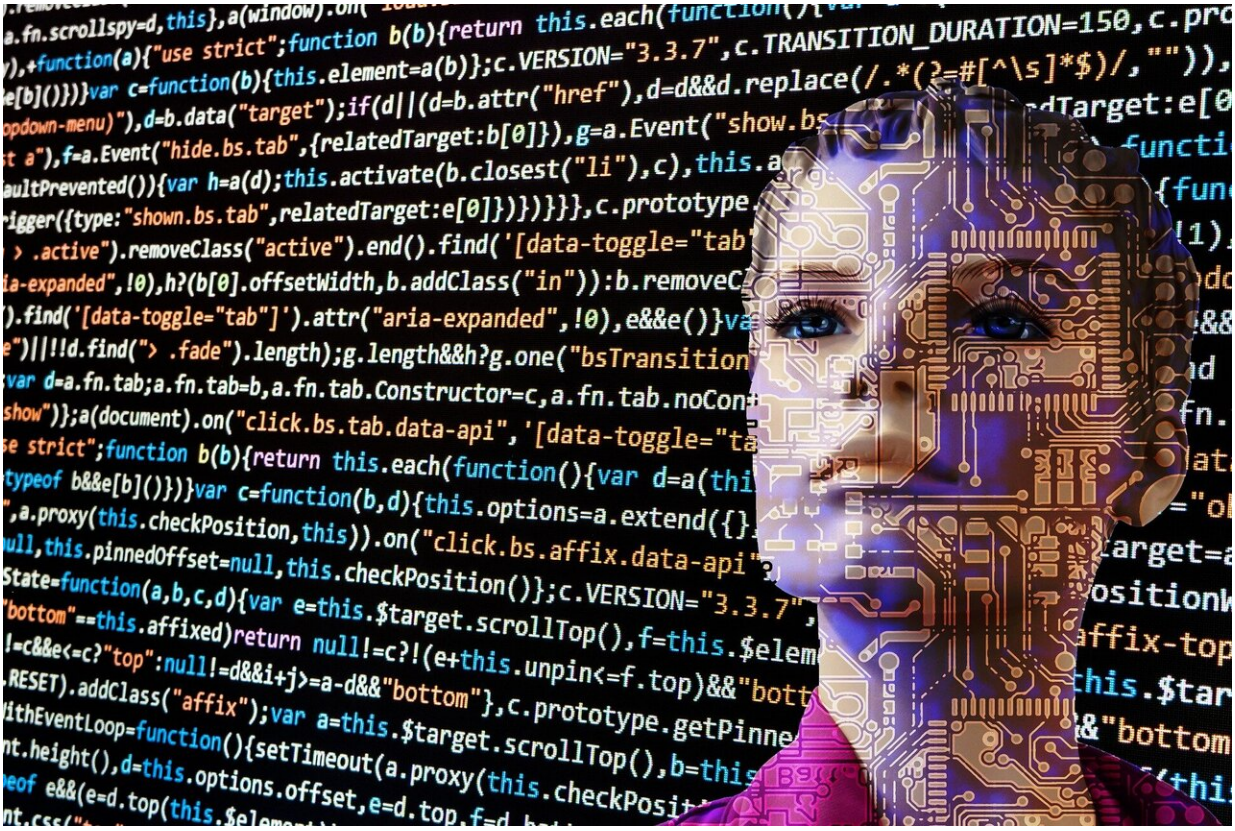
Australia plans to regulate 'high-risk' AI. Here's how to do that successfully

January 18 2024, by Lisa M. Given



Credit: CC0 Public Domain

This week, federal Minister for Industry and Science Ed Husic announced the Australian government's response to the Safe and Responsible AI in Australia consultation.

The response addresses feedback from last year's consultation on artificial intelligence (AI). It received more than 500 submissions, noting "excitement for the opportunities" of AI tools, but also raising concerns about potential risks and Australians' expectations for "regulatory safeguards to prevent harms."

Instead of enacting a single AI regulatory law like the European Union has done, the Australian government plans to focus on high-risk areas of AI implementation—ones with the greatest potential for harm. This could include examples such as discrimination in the workplace, the justice system, surveillance, or self-driving cars.

The government also plans to create a temporary expert advisory group to support the development of these guardrails.

How will we define 'high-risk' AI?

While this proportional response may be welcomed by some, focusing on high-risk areas with only a temporary advisory body raises significant questions:

- how will high-risk areas be defined—and who makes that decision?
- should low-risk AI applications face similar regulation, when some interventions (such as requiring watermarks for AI-generated content) could broadly combat misinformation?
- without a permanent advisory board, how can organizations anticipate risks for new AI technologies and new applications of AI tools in the future?

Assessing "risk" in using new technologies is not new. We have many existing principles, guidelines, and regulations that can be adapted to address concerns about AI tools.

For example, many Australian sectors are already highly regulated to address safety concerns, such as vehicles and medical devices.

In all research involving people, Australian researchers must comply with national guidelines where risk assessment practices are well defined:

- identifying the risks and who might be at risk of harm;
- assessing the likelihood, severity and magnitude of risk;
- considering strategies to minimize, mitigate, and/or manage risks;
- identifying potential benefits, and who may benefit; and
- weighing the risks and determining whether the risks are justified by potential benefits.

This risk assessment is done before research being done, with significant review and oversight by Human Research Ethics Committees. A similar approach could be used for AI risk assessment.

AI is already in our lives

One significant problem with AI regulation is that many tools are already used in Australian homes and workplaces, but without regulatory guardrails to manage risks.

A recent YouGov report found 90% of Australian workers used AI tools for daily tasks, despite serious limitations and flaws. AI tools can "hallucinate" and present fake information to users. The lack of transparency about training data raises concerns about bias and copyright infringement.

Consumers and organizations need guidance on appropriate adoption of AI tools to manage risks, but many uses are outside "high risk" areas.

Defining "high risk" settings is challenging. The concept of "risk" sits on a spectrum and is not absolute. Risk is not determined by a tool itself, or the setting where it is used. Risk arises from contextual factors that create potential for harm.

For example, while knitting needles pose little risk in everyday life, knitters are cautioned against carrying metal needles on airplanes. Airport security views these as "dangerous" tools and restricts their use in this setting to prevent harm.

To identify "high risk" settings we must understand how AI tools work. Knowing AI tools can lead to gender discrimination in hiring practices means all organizations must manage risk in recruitment. Not understanding the limitations of AI, like the American lawyer who trusted fake case law generated by ChatGPT, highlights the risk of human error in AI tool use.

Risks posed by people and organizations in using AI tools must be managed alongside risks posed by the technology itself.

Who will advise the government?

The government notes in its response that the expert advisory body on AI risks will need "diverse membership and expertise from across industry, academia, civil society and the legal profession."

Within industry, membership should include various sectors (such as health care, banking, law enforcement) with representation from large organizations and small-to-medium enterprises.

Within academia, membership should include not just AI computing experts, but also social scientists with expertise in consumer and organizational behavior. They can advise on risk analysis, ethics, and

what people worry about when it comes to adopting new technology, including misinformation, trust and privacy concerns.

The government must also decide how to manage potential future AI risks. A permanent advisory body could manage risks for future technologies and for new uses of existing tools.

Such a body could also advise consumers and workplaces on AI applications at lower levels of risk, particularly where limited or no regulations are in place.

Misinformation is one key area where the limitations of AI tools are known, requiring people to have strong critical thinking and information literacy skills. For example, requiring transparency in the use of AI-generated images can ensure consumers are not misled.

Yet the government's current focus for transparency is limited to "high-risk" settings. This is a start, but more advice—and more regulation—will be needed.

Provided by The Conversation

Citation: Australia plans to regulate 'high-risk' AI. Here's how to do that successfully (2024, January 18) retrieved 28 April 2024 from https://techxplore.com/news/2024-01-australia-high-ai-successfully.html