

AI experts cautiously optimistic about Government's AI Safety Institute announcement

Experts welcome long-awaited Safety Institute as a "great start," but warn that leadership and mission focus are critical to success.

CANBERRA, AUSTRALIA, November 24, 2025 /EINPresswire.com/ -- Today, the Australian Government announced it will establish an Australian AI Safety Institute (AISI), delivering on its commitments to AI safety and responding to calls from Australia's AI safety community dating back to 2023.



Australians For AI Safety

Australians for AI Safety logo

"An AISI is one of the key measures experts have been calling for," said Greg Sadler, CEO of Good Ancestors and spokesperson for [Australians for AI Safety](#). "How Australia navigates frontier AI will determine our prosperity and security. A world-leading AI Safety Institute will give Australia the technical expertise to understand advanced AI, contribute to preventing serious risks, and put us on the path to global leadership."

Australia will join the UK, US, Canada, Japan, South Korea and a growing list of other nations tackling risks from frontier AI through the International Network of AI Safety Institutes.

In March, hundreds of Australian AI experts, public figures and concerned citizens called on the Government to create an AISI, alongside other measures, including guardrails for high-risk AI.

"The announcement of an AI Safety Institute at the start of AI Week 2025 is excellent news," said Jisoo Kim, Director of ClearAI. "Obviously, we're looking forward to more details about the Government's approach to AI safety, governance and the National AI Capability Plan, but this is a great start."

An AISI has broad support beyond AI experts. The Business Council of Australia, the Tech Policy Design Institute, Electronic Frontiers Australia and youth advocacy groups, including Prevention United, have all called for its creation.

The announcement also responds to public demand. This month, a [University of Queensland survey](#) found that fewer than one in four Australians (23%) trust technology companies to ensure AI is safe, suggesting a need for government intervention. 90% of respondents said they would trust AI more if Australia had an AISI working to understand AI risks.

"The creation of an AISI is an excellent move, but we'll be watching the details closely," said Associate Professor Michael Noetel from the University of Queensland, who led the survey. "One of the most important things to get right is leadership. Other parts of Government are focused on current AI risks and on driving AI adoption. The AISI needs to be led by someone focused on the safety of frontier AI development and with credibility in Silicon Valley. They'll need to attract and retain talent and negotiate with AI companies for access and transparency."

"Mission and funding are also important," added Dr Alexander Saeri, AI governance researcher at The University of Queensland and Director of the AI Risk Initiative at MIT FutureTech. "An AISI needs the funding necessary to attract in-demand talent as well as access the computing resources necessary to probe frontier AI systems. We also need a clear mission statement that centres the organisation's focus on frontier AI risks. AI presents many challenges, all of which should be addressed, but an AISI must focus on the risks that are likely to cause severe or even catastrophic harm."

The UK's AISI has a budget of £66 million per year and employs over 100 technical staff.

"Everyone I've spoken to this morning is incredibly excited by this announcement," Sadler added. "This wouldn't have happened without sustained effort from the hundreds of AI experts, researchers and concerned citizens who signed letters, made submissions and contacted MPs. We're grateful to Government for listening to experts and the public, understanding the risks that Australians face from frontier AI, and taking concrete and credible action to keep Australians safe."

Leading AI researchers estimate a 10–50% probability of catastrophic outcomes as AI continues to advance. Last month, over [126,000 signatories](#)—including the world's most prominent AI researchers and notable national security leaders—called for the development of superintelligence to be prohibited until there is a broad scientific consensus that it can be done safely and controllably, with strong public buy-in.

Anthropic, OpenAI and Google have all reported this year that their latest models crossed critical thresholds for assisting with chemical, biological, radiological and nuclear weapons development. Earlier this month, Anthropic disclosed that China-linked hackers had used its AI models in the first documented case of AI-enabled espionage.

"We're seeing frontier AI cross dangerous capability thresholds in real time. An AISI will give Australia the technical expertise to evaluate these systems independently, advise Government on where the red lines should be and how they can respond," said Kim. "An AISI also provides a

formal avenue to work with allies and partners globally. Australian businesses will be able to benefit from knowledge exchanges and dialogue between other Institutes regarding critical AI safety and security issues.”

Dr Toby Ord, author of *The Precipice: Existential Risk and the Future of Humanity* and Senior Research Fellow at Oxford University, said, “An Australian AI Safety Institute would allow Australia to participate on the world stage in guiding this critical technology that affects us all.”

Australians for AI Safety will continue to advocate for the appropriate regulation of high-risk AI, and for Australia to lead on international efforts to ensure advanced AI is developed safely.

Mr Gregory Sadler

Good Ancestors

+61 401 534 879

[email us here](#)

Visit us on social media:

[LinkedIn](#)

[Other](#)

This press release can be viewed online at: <https://www.einpresswire.com/article/869773624>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2025 Newsmatics Inc. All Right Reserved.