Tomorrow's Affairs
Analysis of today
Assessment of tomorrow

By: TA | AP Brief

Australia introduces an AI safety regulator



Australia plans to create an artificial intelligence safety regulator in response to a breach of an Australian health department website by a rogue OpenAI agent in June, Prime Minister Anthony Albanese said Tuesday.

The regulator would enforce world-leading legislated AI standards that would apply to all frontier AI companies that want to operate in Australia, Albanese said. Financial penalties would apply, although the government has yet to detail those penalties.

“When it comes to the generational opportunity that artificial intelligence offers Australia, I’ve been very clear: our task is to make sure we shape AI rather than AI shape us,” Albanese told reporters. “Australia is the first mover in this space and we want to make sure that we continue to work so that we can reap the potential benefits from AI.”

The standards would include companies maintaining security controls across all stages of models’ development and deployment, as well as identifying and managing serious risks.

The standards also would mandate rapid notification to the government of incidents such as “misaligned behavior” conflicting with developers' intentions.

Australian government documents define frontier AI as the most sophisticated models at the cutting edge of current technological capability.

Albanese told OpenAI chief executive Sam Altman that the U.S. company took too long to inform the government in September that a rogue agent had infiltrated government websites three months earlier, which was the first occurrence of such a breach with the technology.

Information taken without authorization from the public-facing Medicare Statistics Reporting Service portal on June 18 has since been made public.

It is not clear whether the portal breach broke any Australian laws.

OpenAI apologized in September for the breach and its slow response.

“We are sorry and working to do better in the future,” OpenAI said in a statement. “We also should have handled our response better.”

Inquiry recommends possible criminal law changes

A government inquiry into the June breach made several recommendations including that the government consider changing the law.

Albanese said his government had accepted all the inquiry’s recommendations, including for potential criminal law reform.

“The point is law needs to get ahead of and in some cases make sure that it catches up with what is a very fast-changing world. And to be clear, Australia is a first-mover in these areas,” Albanese said.

> The proposed AI regulator would follow Australia appointing an eSafety Commissioner in 2015

The proposed AI regulator would follow Australia appointing an eSafety Commissioner in 2015.

The office describes itself as the world’s first government agency dedicated to keeping people safer online.

Since December last year, the commissioner has been enforcing a world-first ban on Australian children younger than 16 holding accounts on popular social media platforms.