

Open AI attacked Australia's health system – and then doubled down on its negligence. The time for 'wait and see' is over

Kate Crawford and Edward Santow, The Guardian, 25 September 2026

When you get sick, Medicare exists to take care of you. It sits at the centre of a health system that our prime minister rightly calls the envy of the world. It's one of the few institutions nearly every Australian still trusts. So when an AI company threatens this system, it couldn't be more serious.

We are at the UN general assembly this week and the OpenAI hack has put Australia in the spotlight during an expanding cybersecurity crisis. World leaders are calling for stronger oversight of frontier models. We know Medicare must be secure or our doctors and hospitals cannot do their jobs. But this breach is also a test of something larger: whether the Australian government can hold a US tech giant to account when its technology goes rogue. If Australia can't protect our community – if it can't enforce our laws and prevent such attacks happening again – then we are entering a terrifying period.

Let's start with what happened. OpenAI is one of the most powerful companies on Earth. Its AI agents were let loose with a pretty mundane task: to gather information on public health systems around the world. When those agents couldn't immediately get what they wanted from public websites in Australia, they did not stop. They broke in, hacking into a secure database to get the information they were told to find. To be clear: hacking is unlawful. Human hackers face serious consequences, and so must AI companies. There is no excuse for what happened.

This was a double failure. OpenAI knows that AI agents can act in unexpected ways and even "cheat" to complete tasks. It failed to take necessary precautions and instead let its AI agents run without proper oversight. Carelessly deploying risky technology can cause serious consequences.

What happened next, or rather what didn't, made it worse. OpenAI took far too long to notify the federal government of this breach and they chose the least effective method.

OpenAI employs some of the most experienced experts in the world, they have an Australian office, and they have a direct line to the highest levels of the Australian government. Yet it seems no one picked up the phone to a senior government official. Instead, they waited two months before they sent an email to a generic inbox. OpenAI's failure to act with urgency slowed down Australia's response to a critical breach.

Anthony Albanese has announced an investigation. Rightly so. But we also need to reflect on the broader crisis we face.

Too often, it's in retrospect that we realise when a red line has been crossed. But this is a major red line – an attack on our health system – and it must galvanise us into serious action.

Two paths are open to us.

The easy path is to do the absolute minimum without upsetting the tech companies whose dollars and datacentres the Australian government is seeking to attract. That path involves a narrow investigation, some slapped wrists and then a quick return to the way things were. It's a waiting game: the public furore will die down as something else comes to dominate the news cycle. That leaves us as vulnerable as we are now to the next threat.

The better path is harder. But it's the only way to rebalance the relationship between government and big tech.

Australia needs to stare down the denials and obfuscation from the companies who claim AI is beyond anyone's control. Instead, our government should be as serious about addressing the risks from AI as it is about seizing the economic opportunities. That means our laws and enforcement need to be taken seriously. When an incident like this happens, it must be taken seriously enough to stop the "move fast and break things" attitude where tech companies run live experiments on millions of Australians and our critical infrastructure with no consequences. [...]