AI tech pioneers' fears of AI beyond human control show it is past time for an effective regulator

Melissa Donnelly, ACTU Secretary

One of the hardest challenges of Artificial Intelligence is to sort fact from fiction. It is tempting to dismiss the weekend proclamations from AI company CEOs - Dario Amodei, Sam Altman and Elon Musk - that the powerful technology they're developing is advancing too quickly because their warnings are motivated by greed and self-interest. Obviously, they have a lot of money riding on people believing that.

But the resignation of AI researchers, voicing genuine concerns about the fate of humanity in the next few years if we don't slow down the rate of progress, shouldn't be ignored.

We now have rare agreement between rivals Anthropic and OpenAI that their models should be developed at a more balanced rate, with time for outside evaluation, to secure their safety.

The founders and people building this technology openly express their fears that AI, beyond the control of humans, is absolutely possible in the near future.
You don't need to believe the most cataclysmic future scenarios to recognise that the capabilities of the most advanced AI models pose serious risks right now.

The warnings are clear, and so is the lesson: Australia must not wait any longer to regulate this all-powerful new technology. This may be our last chance to set the trajectory of AI's development on our own terms, to protect our rights, our values and our fair go.

Unions are deeply concerned about working peoples' rights in the age of AI. Much of our focus is on the workplace but working people's lives don't end when their shift does. Working people, and all Australians, deserve basic protections from these new technologies as they threaten the rapid transformation of all facets of our lives.

As it stands, we are allowing large tech companies - whose primary motivation is selling products and making money – to determine how this technology is being developed and in whose interests.

We have already seen that there are clear risks in ungoverned and poorly motivated AI model development and major risks in the misuse of this technology. Some of these risks are in the here and now: the use of AI to create child sexual abuse material or sexually harass people, the spread of misinformation and disinformation in our democracy, and the use of models to enable cyberattacks on critical infrastructure and important businesses.

While it is encouraging to see a company like Anthropic publish its efforts to disrupt malicious use, like the potential development of bioweapons, we can't rely on privatising the regulation of these models.

OpenAI's recent failure to govern a swarm of agents resulted in the much-publicised Hugging Face attack. How hundreds of agents were able to operate undetected and maintain an unsanctioned communication channel for days raises serious questions about the company's safety protocols and governance.

Australia cannot trust our fate to the decisions of these companies. We have a right and a responsibility to take sensible measures to protect ourselves.

This is why the union movement has been calling for a licensing regime for AI models, including proactive risk mitigation requirements for developers and deployers of AI in Australia.

There are serious risks already in evidence and Australians deserve to know that the developers of the technology have, at a minimum, taken steps to mitigate those risks. And if they haven't, they shouldn't be permitted to profit from Australians.

A licensing regime would mandate ongoing transparency over AI products for sale in this country, including regular auditing of their safety, the development of guardrails against high-risk use, effective governance, and mandatory reporting of major incidents.

The natural place for this licensing regime to be administered is in a beefed-up Artificial Intelligence Safety Institute, created by the Albanese Government earlier this year. The Institute will need greater powers to face the scale of the existential threat the founders and builders of this tech are now sounding the alarm on.

For any licensing regime to be effective, non-compliance should carry stiff penalties and ultimately the ability to ban these models from access to the Australian market if they cause harm by operating unsafely.

Australia is a world-leader in effective regulation of digital platforms. The success of the social-media ban for under 16s has led other countries to consider the same approach to limit the harms of that technology. The rapid pace of AI development though means we can't afford to wait two more years, let alone twenty years before setting up a licencing regime.

Big business detractors of effective regulation point to the US and claim that we should let it rip too, but that comparison has aged like milk. Even setting aside how various US State Legislatures are regulating the technology, and despite Donald Trump dismissing these concerns as a hoax, the US President has created a de facto regulator by acting as an approver - or not - of new AI model releases. Of course, he only considers what's in America's interest.

Rather than being at the whim of whichever tech titan finds favour in the Oval Office, we have an opportunity to work with peer democracies on the effective regulation of AI. But it requires someone to move first. Australia should do it.

Workers are already victims of the misuse of AI. Workers have experienced harassment, deepfakes, theft of their creative and academic content, and harmful intrusive surveillance. Regulating this technology to enforce decent safety standards will protect workers and make our communities safer. If the tech pioneers are even partly right, there's no time to waste to prevent advanced AI systems moving beyond all our control.