



Guardrails for the development of Artificial Intelligence

June 2025

“You seek for knowledge and wisdom, as I once did; and I ardently hope that the gratification of your wishes may not be a serpent to sting you, as mine has been.”

Dr Victor Frankenstein in Mary Shelley’s novel *Frankenstein, or The Modern Prometheus* (1818), referencing the serpent in the Garden of Eden in Genesis

The February 2021 Synod meeting resolved:

The gospel calls us to relate to each other with love, treating each other with dignity and respect, and to condemn exploitation and abuse of vulnerable people. God’s people are called to pursue justice, including by empowering those who are exploited and abused....

In our age, there is a need to prevent and address human rights abuses online, including acting against the promotion and facilitation of child sexual abuse.

Artificial intelligence (AI) is rapidly developing. While AI can have significant benefits for some uses, unfortunately, there are many risks and harms that AI can also facilitate or create.

In October 2024, we issued an action outlining how AI is facilitating child sexual abuse, primarily through the creation of fake sexual images of children to be used for sexual extortion of the child. Such extortion can include coercing the child to provide real images of themselves.

Other risks from the development of AI technology include:

- Undermining democracy by being used to generate disinformation, such as fake videos of politicians that make it impossible to know reality from propaganda;
- Gathering data about people to manipulate them to a variety of ends, including the purchase of harmful products;
- The creation of AI weapons for military use;
- The ability to manipulate us to harm each other; and,
- Workplace surveillance used to exploit workers, including imposing unreasonable work demands and collecting and selling of data about workers without their informed consent.

In December 2024, Professor Geoffrey Hinton, regarded as the “godfather” of AI and a former employee of Google, publicly stated that the pace of development of AI carried a 10% to 20% risk that it would lead to human extinction in the next 30 years.¹ He has stated that technology corporations cannot be trusted to develop AI safely. “So just leaving it to the profit motive of large companies is not going to be sufficient to make sure they develop it safely”, he said. “The only thing that can force those big companies to do more research on safety is government regulation.”²

While such an assertion may seem like a wild prediction, the Oxford University Future of Humanity Institute (which closed in 2024) and the Cambridge University Centre for the Study of Existential Risk both considered that the development of AI poses an existential risk to humanity. Toby Ord from the Future of Humanity Institute published a book in 2020, *The Precipice*, which assessed the possible existential threats humanity faces in the next century. He placed risk assessments on these threats to the future of humanity, with the leading threats being:

- 1 in 10 from the development of AI that is smarter than humanity;
- 1 in 30 from the development of a man-made pandemic;
- 1 in 1,000 from human-induced climate change; and,
- 1 in 1,000 from nuclear war.

¹ Dan Milmo, “‘Godfather of AI’ shortens the odds of the technology wiping out humanity over next 30 years’, *The Guardian*, 28 December 2024.

² *Ibid.*



A 2016 survey of AI researchers found that 48% believed there needed to be an increase in AI safety research. Only 12% believed there should be less research into AI safety.³ So far, it has been impossible to encode human values in AI to ensure it always acts in our best interests.⁴ It would be best to take regulatory precautions to ensure that such predictions remain the stuff of science fiction.

On his first day in office, President Donald Trump revoked a 2023 executive order from President Biden that put guardrails on the development of AI, including concerning public health and the economy.

Example of AI increasing people killing each other⁵

Meta, the owner of Facebook, had developed AI algorithms (computer programs) designed to maximise user engagement on Facebook. The algorithm decided that getting people outraged got them to stay on Facebook longer, increasing Facebook's profits from advertisers.

In Myanmar, Facebook has been the primary news source for millions of people and the most important platform for political mobilisation. So, in 2016 to 2017, when there was a campaign by the Myanmar military and Buddhist extremists to exterminate people from the Rohingya ethnic minority in Myanmar, the Facebook algorithm promoted extremist propaganda encouraging the mass murder of Rohingya. Between 7,000 and 25,000 Rohingya were slaughtered, and around one million were forced to flee the country.

In 2018, a UN fact-finding mission concluded that the Facebook algorithm had played a "determining role" in the mass murder campaign.

The Australian Government has not decided how it will regulate the development of AI. Those seeking to profit from AI have been advocating for the Australian Government not to get in their way with measures to protect the community from the harms their products might cause. We need to make sure community voices are heard about the importance of proper regulation.

What You Can Do

Write polite and respectful letters to:

The Hon Anthony Albanese MP
Prime Minister
PO Box 6022
House of Representatives
Parliament House
Canberra ACT 2600

Senator The Hon Tim Ayres MP
Minister for Communications
PO Box 6100
The Senate
Parliament House
Canberra, ACT, 2600

Salutation: Dear Prime Minister

Salutation: Dear Minister

Points to make in your letter:

- Express deep concern at the risks that AI can pose to our future well-being.
- Urge that the Australian Government put in place robust safeguards in law over the development of AI in Australia to reduce the risks of misuse of AI that will cause serious harm.
- Urge the Australian Government to be active globally in seeking international agreements between governments to ensure the safe development of any AI technologies that would otherwise potentially cause substantial harm.

Please send us copies of any replies you get to your letters.

³ Toby Ord, 'The Precipice', (London: Bloomsbury Publishing, 2020), 151.

⁴ Ibid., 144.

⁵ Yuval Noah Harari, 'Nexus. A Brief History of Information Networks from the Stone Age to AI', (New York: Penguin Random House, 2024), 195-200.