

Technology [AI](#)

OPINION

The AI horse has bolted. It's time for the nuclear option



[Alan Finkel](#)

Former chief scientist

August 7, 2023 — 3.30pm

A biosecurity expert at the Massachusetts Institute of Technology recently challenged his students to use publicly available generative artificial intelligence (AI) tools to identify a virus capable of unleashing a pandemic. [It took the students only an hour](#) to produce a list of four candidate viruses.

In contrast to AI's potential to contribute to society, new concerns are published almost daily about its potential to be used to design powerful chemical toxins, to replace actors with their likenesses without their consent, to perpetrate undetectable scams, spread misinformation during elections and build autonomous weapons.



“Because it is hard to regulate AI, should we do nothing? No.” Former Australian chief scientist Alan Finkel. EAMON GALLAGHER

Cries for regulation are issued almost daily. [The European parliament has approved sweeping draft laws](#) which, if legislated, will restrict the use of AI and mandate the role of human oversight, data quality and accountability. European companies are pushing back, concerned that the proposed legislation will not be effective at tackling AI risks and will damage the competitiveness of European businesses.

The trouble for regulators is that the horse has bolted. AI is everywhere. Regulations will be adhered to by high-integrity organisations and they will dutifully label their products. However, the rules will inevitably be ignored by criminal organisations, terrorist groups and rogue states.

To add to the dilemma, if some countries were to apply the brakes on development, others with a history of supporting terrorism and developing nuclear weapons would ignore the call and leap ahead in the AI race.

Because it is hard to regulate AI, should we do nothing? No. The risks are so great that burying our collective heads in the sand is not the answer. Calls by [technology leaders for a hiatus in AI development](#) are not the answer. Regulations that only affect the well-intentioned, high-integrity developers and users of AI are not the answer.

Perhaps the answer will come from thinking laterally. What can we learn from the nuclear non-proliferation treaty that entered into effect in 1970?

While not perfect, it slowed the spread of nuclear weapons, and arsenals today are about one-fifth of what they were 50 years ago. What made the treaty possible was that the cost and scale of developing nuclear weapons is so large that the regulated targets like the silos, reactors and enrichment facilities can be monitored by international audit teams and by satellite

surveillance to verify compliance. The problem is not solved, but we have enjoyed decades of respite.

The equivalent targets to regulate in the world of AI technology are the facilities that produce the microchips that are essential to the operation of AI algorithms. This is well recognised by the US government, which has imposed punitive restrictions on the export to China not only of advanced microchips, but also the production equipment to make them.

While it is well known that the dimensions and cost of transistors on a microchip have been shrinking for decades, what is less well known is that the dimensions and cost of fabrication plants to make the increasingly powerful microchips have been inexorably rising, with the cost of building a modern fabrication plant estimated to be at least \$US3 billion and as much as \$US20 billion. These massive and complex plants are huge and could be easily monitored by international audit teams.

To be successful, an AI non-proliferation treaty would have to be agreed by every technologically advanced country or jurisdiction, including the US, China, Europe, Taiwan and many others, despite the certain objections of the defence communities in every country. Exceptions could not be allowed because non-participation would be an enormous advantage in the technological arms race.

One of the most important details of the treaty would be agreement on the level of capability that should not be exceeded. The definition would have to be stunningly simple so that it could not be ambiguously interpreted. For example, the level of capability could be defined as the maximum number of transistors per microchip. Limits such as no more than 100 billion might be agreed.

Of course, we need computer microchips for our smartphones, our banking systems and our satellite navigation systems. But do we need the extraordinarily powerful microchips that power generative AI? What defines the line between adequate power and ultimate power? What would we be giving up in exchange for protection against malevolent uses of AI?



Is it time for the nuclear option? GETTY/SUPPLIED

Like the effort to tackle climate change, negotiating an international non-proliferation agreement would require collaboration on an exceptional scale. Given the explosive rate of development in AI, the global community cannot afford to drag its feet. For internationally agreed approaches to be efficacious before the AI tsunami doubles or quadruples again, the aim would be to negotiate an agreement within a year or two. In acknowledgment of the newness of the approach, the agreement would trigger its own renegotiation after an interval of five years or thereabouts.

Who will show leadership on negotiating an AI non-proliferation treaty? It is a collective responsibility and certainly one to which Australia could contribute.

Alan Finkel is an Australian neuroscientist, inventor, researcher, entrepreneur, educator, policy advisor, and philanthropist. He was Australia's chief scientist from 2016 to 2020.

The Opinion newsletter is a weekly wrap of views that will challenge, champion and inform your own. [Sign up here.](#)



Alan Finkel is an Australian neuroscientist, inventor, researcher, entrepreneur, educator, policy advisor, and philanthropist. He was Australia's chief scientist from 2016 to 2020.
