

RSIS Commentary is a platform to provide timely and, where appropriate, policy-relevant commentary and analysis of topical and contemporary issues. The authors' views are their own and do not represent the official position of the S. Rajaratnam School of International Studies (RSIS), NTU. These commentaries may be reproduced with prior permission from RSIS and due credit to the author(s) and RSIS. Please email to Editor RSIS Commentary at RSISPublications@ntu.edu.sg.

Artificial Intelligence Governance: Lessons from Decades of Nuclear Regulation

By Karryl Kim Sagun-Trajano

SYNOPSIS

To some experts and policy makers, the existential threat posed by nuclear weapons has just been paralleled by artificial intelligence. Numerous calls for an international body that will govern the development of artificial intelligence technology have been made, citing the International Atomic Energy Agency as a model. While it is not a foolproof one-to-one comparison, those working on regulation of artificial intelligence can draw useful lessons from the nuclear experience, including how to balance governance and peaceful uses of the technology.

COMMENTARY

For decades, the existential threat posed by nuclear weapons has not been paralleled, that is, until the advent of artificial intelligence (AI) in recent years. In fact, recent discussions on AI often compare it with nuclear weapons.

In May 2023, the [Centre for AI Safety](#), a US-based research and field-building non-profit, released a statement declaring that the existential threat from AI must be prioritised together with nuclear war. It also cited both as societal-scale risks. Signatories included AI scientists such as [Geoffrey Hinton](#) and [Yoshua Bengio](#) (two of the three so-called AI godfathers), CEOs of prominent tech companies, professors from top universities, among many other notable figures.

There have also been calls for an international AI agency equivalent to the International Atomic Energy Agency (IAEA) to govern AI development and application. While on a global tour in June 2023, [OpenAI CEO Sam Altman](#) rallied support for AI governance, making a reference to the IAEA as a model for placing guardrails on a “very dangerous technology”. Shortly after, [UN Secretary-General António Guterres](#) supported the creation of an AI watchdog, which would act like the IAEA.

While it is not an absolute one-to-one comparison, those working on AI regulation have much to learn from the nuclear experience.

A Reactive Regulation?

In the case of nuclear regulation, the call for it was made by physicist Alvin M. Weinberg before the US Senate's Special Committee on Atomic Energy in December 1945 – four months after the atomic bombing of Hiroshima and Nagasaki. It is easy to be critical about this reactive approach to nuclear regulation. To be fair however, nuclear development then was kept under wraps as the objective was to produce a new weapon of war – one that was unprecedented in its potential for destruction.

Such ironclad secrecy is not needed in the case of AI. As this is a general-use technology, the likelihood of it harming humanity has been discussed, not just in closed-door meetings, but also in accessible public discourse.

Not everyone has access to civilian uses of nuclear technologies, in contrast to how anyone can utilise various AI tools. In the [1940s](#), the US government controlled nuclear technologies tightly and preserved them almost exclusively for military use. It changed policy in the [1950s](#) when it realised that continued curtailment of commercial and private uses of nuclear technologies will render them laggards in the nuclear race.

The IAEA was officially established under the United Nations in July 1957, with a [statute](#) approved by 81 nations in October 1956. This was a significant milestone. If done right, the process of establishing AI regulation need not go through a similar time-consuming back and forth track that spanned decades.

Lessons from a Mature Technology and Regulating Agency

There are lessons which AI regulators can learn from the well-experienced IAEA and its regulation of nuclear technologies as the latter has well-established and tried and tested elements of good technological governance.

First, a unifying body for the peaceful uses of AI, such as the proposed “International Artificial Intelligence Agency (IAIA)”, ideally under the United Nations as well, could help to align the divergent [US-EU](#) regulatory plans on AI, [China's](#) rules on AI use, [ASEAN's](#) regional guidelines on AI, and the intended policies of other states. Internationally beneficial norms could be established to harness the potential of AI while mitigating the risks. Time is of the essence as the number of cases of [AI misuse is rising](#) at an alarming rate. This agency can take after the [IAEA](#), which stands as the world's “Atoms for Peace” organisation, with member states and global partners promoting safe, secure, and peaceful nuclear technologies.

Second, information sharing among stakeholders could benefit AI governance. In AI research, industry has [surpassed](#) academia in the production of AI systems. This ascendancy of industry is not necessarily a bad thing as the future of AI regulation may largely depend on an interplay of state laws and industry standards. The proposed IAIA could aid in melding these well. The nuclear experience makes an excellent case for information sharing among various stakeholders, as the IAEA's

stance is that frequent [information exchange](#) is vital for emergency preparedness and response. It provides not just member but also public access to information on power reactors, nuclear data, and research reactors.

Third, capacity building activities could also reduce the [AI divide](#) between more technologically advanced countries and those lagging behind. This would help to curtail the risk of AI feeding global inequality, amongst other things. Within the IAEA, there are efforts to [build capacity for nuclear safety](#). These stand on four pillars: (1) education and training, (2) human resource development, (3) knowledge management, and (4) knowledge networks. These pillars could well be transposed to (or at least be a springboard for) AI.

Challenges to AI Governance

Unfortunately, the threat of AI, relative to nuclear weapons, is stealthier. Any ordinary looking office space can be a hub for the criminal use of AI. It would be quite inconspicuous for a rogue state to trigger an anonymous AI-powered attack, whereas one with a nuclear-weapons capability could not even threaten to use a nuclear weapon at another state without causing an international crisis. Nuclear meltdown accidents are likewise more tangible. Furthermore, nuclear power plants are massive constructions, and smaller reactors, while in existence, are generally still experimental.

Conclusion

Although IAEA regulation is far from perfect, it is important for those working on AI governance to watch and learn from the factors influencing nuclear governance. In spite of the IAEA, some states have the tendency to be extremely cautious and thus overly govern or politicise the technology.

The risks of nuclear technologies and AI have been highlighted by experts, but there are significant differences in their governance. These lie partly with the maturity of nuclear as a technology. Nuclear technologies had a long, painful, and tortuous path to regulation; all while accommodating the inalienable right of states to research, produce, and use them for peaceful purposes.

The regulators of AI cannot hinge entirely on those involved in nuclear governance. They will have their share of achievements and mistakes. They have much to learn from the nuclear experience, including how to balance governance and peaceful uses of the technology. Over the years, nuclear technologies have grappled with, and, in some significant ways, managed to advance within a tightly regulated regime and close international scrutiny. The prospect for the positive development and good governance of AI technology is not an illusion.

Karryl Kim Sagun-Trajano is a Research Fellow at Future Issues and Technology (FIT), S. Rajaratnam School of International Studies (RSIS), Nanyang Technological University (NTU), Singapore.
