

www.rsis.edu.sg

No. 017 – 30 January 2024

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.

Singapore's Proposal on Global Generative Al Governance

By Jose Miguelito Enriquez

SYNOPSIS

Singapore's proposed Model AI Governance Framework for Generative AI is a step in the right direction for global generative AI governance, but global regulation advocates face a tough road ahead. It is necessary to engage all parties involved to reach an equitable and viable structure going forward.

COMMENTARY

On 16 January 2024, the AI Verify Foundation and the Infocomm Media Development Authority (IMDA) published their <u>Proposed Model AI Governance Framework for</u> <u>Generative AI</u>. While it is the third iteration of their Model AI Governance Framework, this version is the first to focus squarely on regulating Generative Artificial Intelligence (Gen AI) models such as Google's Gemini and OpenAI's Generative Pre-Trained Transformer (GPT).

The rapid mainstreaming of Gen AI models, spurred by the launch of <u>OpenAI's</u> <u>ChatGPT chatbot</u> in November 2022, is responsible for the so-called "AI boom". The swift development of Gen AI models consequently led to an increasingly urgent need to regulate AI and guarantee its secure development and use.

Several regions have already responded to this need. Last December 2023, the <u>European Union</u> reached a provisional deal on its AI Act. In October 2023, <u>US</u> <u>President Joe Biden</u> signed an executive order to ensure safe AI development in the United States. Still, other countries, like the <u>United Kingdom</u>, have put a hold on their plans to pass legislation out of concerns it could restrict innovation.

A Blueprint for Global Dialogue

Singapore enters the conversation on Gen AI governance not with a domestic law, but a guiding framework towards a global AI regulatory system by focusing on nine key elements to build confidence in the AI ecosystem: accountability, data, trusted development, incident reporting, testing and assurance, security, content provenance, safety and alignment, and ensuring AI for the public good.

In this document, Singapore crafts a proposal that creates a trustworthy AI ecosystem for consumers but also an environment conducive to innovations from AI developers and related businesses. By providing a holistic discussion on Gen AI governance, the Model Framework is a useful blueprint for global conversation on AI governance issues.

The Model Framework provides concrete policy recommendations by drawing parallels to other industry regulations, such as shared accountability between AI model developers and AI-based application developers patterned after responsibility models in the cloud computing industry.

The document also clearly states which existing legal statutes need to be updated to cater to the novel use cases caused by Gen AI, such as in product liability protections and personal data protection. Amending data protection statutes has become salient as the training data used by AI developers, a <u>once overlooked issue</u>, is now subject to close scrutiny.

Finally, the Model Framework also explores an issue of AI use that is currently overlooked – its sustainability. While it is currently difficult to pin down the exact environmental impact of AI, current estimates show that Google's AI operations alone could produce a carbon footprint <u>similar to that of a small country</u>.

Developers contend that the current environmental impact of AI is overstated and that servers used for AI operations consume considerably less electricity than traditional data centres. However, a recent study estimated that by 2027, AI servers manufactured by chipmaker Nvidia are projected to consume <u>134 Terawatt hours</u> (<u>TWh</u>) of power. This is comparable to the consumption of the <u>Bitcoin mining network</u> today.

It is imperative that the environmental costs of AI are regularly monitored. In this regard, the Model Framework's recommendation to build efficient computing centres and incentivise green energy use should be accompanied by strict requirements for AI developers to report their operations' energy consumption and carbon emissions.

Mitigating Harms and Navigating Contentious Issues

The Model Framework also shares proposals to minimise harm in areas where malicious AI use could lead to societal harm, such as in deepfaking. It rightly points out the urgency to institute standardised content provenance labels to make it easier for users to know when an image or video has been edited or wholly generated through Gen AI – a harm that Singapore recently faced when a deepfake video of <u>Prime</u> <u>Minister Lee Hsien Loong</u> surfaced online.

However, provenance labels, such as the watermarks and cryptographic provenance identified in the framework, will only be effective if all stakeholders agree on a single, interoperable, tamperproof labelling standard. While <u>work on open standards is</u> <u>ongoing</u>, a coordinated and sustained dialogue across the public and private sectors on this key issue is needed to keep the momentum and achieve this goal.

Moreover, while the Model Framework maps out an ambitious policy roadmap that tackles the entire AI development process, it appears to be less instructive on managing copyright concerns, a topic that could potentially become the most contentious in Gen AI governance.

The issue recently came to light when <u>several lawsuits alleged</u> that AI developers trained their models based on the copyrighted works of authors, journalists, and musicians without obtaining prior permission.

The Model Framework does not make a concrete proposal to resolve these concerns. It appropriately stated that continuous dialogue is required to produce a viable solution that balances copyright concerns with the need for AI developers to access quality training data.

Elsewhere, countries have also grappled with how to move forward in resolving this issue. In the UK, an early proposal to allow AI developers to freely use copyrighted material as training data was criticised by several members of Parliament. In the US, several lawmakers supported a proposal to require AI companies to pay licensing fees to use copyrighted material but was met with criticism from AI industry executives.

It is still unclear what a viable solution to AI's copyright dilemma would be. However, policymakers around the world need to explore possible options now to keep pace with the innovation taking place within the AI industry. Concerns within the industry must also be tempered with the rights of creative individuals whose livelihoods and body of work are at risk from the continuing intrusions of Gen AI.

The Road to Global Regulation

As the AI boom shows no signs of slowing down, managing Gen AI's most disruptive effects should be a discussion taking place at the international level. Singapore's latest Model AI Governance Framework offers a compelling roadmap to advance a global framework and a state-led response to today's challenges in Generative AI governance.

However, even with elevated enthusiasm for Gen AI governance, it may take a while to arrive at a global agreement. If the <u>experience of the EU with the AI Act</u> is a sign, these negotiations could become very heated and contentious, and at times even break down due to divergent state and stakeholder interests.

To prevent a repeat of the protracted discussions in the EU, advocates for global AI governance like Singapore could benefit from initially convening informal dialogues with a smaller group of like-minded governments as well as with business leaders, civil society organisations, and AI developers.

Continuously engaging in dialogues will help generate cross-stakeholder support around the proposals laid down in the Model Framework, which will then provide momentum once the conversation is expanded to a wider global forum.

Jose Miguelito Enriquez is an Associate Research Fellow in the Centre for Multilateralism Studies at S. Rajaratnam School of International Studies (RSIS), Nanyang Technological University (NTU), Singapore. His research interests include digital economy governance in ASEAN, populist foreign policy, and Philippine politics and foreign policy.

> S. Rajaratnam School of International Studies, NTU Singapore Block S4, Level B3, 50 Nanyang Avenue, Singapore 639798 T: +65 6790 6982 | E: <u>rsispublications@ntu.edu.sg</u> | W: <u>www.rsis.edu.sg</u>