



# Grok – An Emerging AI Governance Moment for Southeast Asia

*Karryl Kim Sagun Trajano*



*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.*

## **Grok – An Emerging AI Governance Moment for Southeast Asia**

*By Karryl Kim Sagun Trajano*

### **SYNOPSIS**

*Indonesia and Malaysia's bans on xAI's Grok mark a regulatory pivot, moving Southeast Asia from late adoption to early action on AI safety. The case provides a leadership opportunity to shape context-specific norms on AI safety, but also risks fragmentation without ASEAN-wide coordination and safeguards.*

### **COMMENTARY**

While much attention has focused on the [row](#) between the United States and the United Kingdom over the latter's banning of [Grok](#) (Elon Musk's AI-powered chatbot on X), Indonesia and Malaysia had already banned the platform days earlier. These actions seem to be establishing a regional pattern, with the Philippines also joining the list of countries banning Grok. This marks an important regulatory pivot: Southeast Asian states are moving from late adopters to early movers on a highly contested frontier of AI safety, online harms, and platform governance.

[Indonesia's](#) decision on 10 January to temporarily block access to Grok marked the first time a state intervened directly against the platform. The move was triggered by concerns over the tool's "digital undressing" capability, which enables the creation of non-consensual, sexualised nude or near-nude deepfake images, including of children.

[Malaysia](#) followed suit within a day, imposing a similar temporary restriction after documenting repeated misuse of the system to produce obscene and manipulated content, despite prior regulatory warnings and safeguard mechanisms that largely depended on post hoc user reports.

The [Philippines](#), announcing an official ban on 15 January, has characterised Grok's "undressing" capability as a cybercrime, placing it within the category of online sexual abuse and exploitation of children.

In these cases, authorities presented the restrictions as conditional and corrective, indicating that access would only be restored once xAI and X demonstrated compliance with domestic legal obligations and implemented more robust, preemptive safety measures.

Crucially, these interventions were based not on moral regulation but on established policy rationales for digital safety, rights protection, and platform accountability, as emphasised by Indonesia's [Communication and Digital Affairs Ministry](#), Malaysia's [Communications and Multimedia Commission](#) and the Philippine's [Department of Information and Communications Technology](#).

### **Beyond Religious Conservatism**

Considering that Indonesia and Malaysia rely heavily on largely Islamic moral frameworks, and the Philippines is predominantly Catholic, a knee-jerk interpretation might attribute the bans to religious conservatism. However, this framing risks overlooking the political and regulatory dynamics at work, especially since other conservative or religious societies have not taken similarly aggressive action against Grok despite encountering comparable online harms.

What differentiates Indonesia, Malaysia and the Philippines is a convergence of political incentives, regulatory experience with platform controls, and global reputational considerations. These governments have previous experience in blocking or restricting platform access over concerns such as [pornography](#), [gambling](#), and [online sales of illegal items](#), and those experiences have provided them with legal and operational tools to act quickly when faced with a new category of harm. In this sense, these measures demonstrate a pragmatic application of existing statutes to emerging technologies rather than being driven solely by cultural or religious sensibilities.

It is also notable that Indonesia, Malaysia, and the Philippines had acted ahead of many Western and more technologically advanced jurisdictions. This development comes at an interesting time, as in December 2025, the United States moved to roll back what it described as "[cumbersome](#)" AI regulation, signalling an even more anti-governance stance. The United Kingdom, meanwhile, [warned](#) that Grok would no longer be permitted to self-regulate, with urgent investigations underway regarding a proposed [ban](#).

But even as the US and European countries continue to deliberate on governance responses, Indonesia, Malaysia and the Philippines have already implemented enforcement-oriented measures to tackle specific harms. In doing so, these three countries have shifted from being regarded as slow adopters to emerging early movers in AI oversight.

## Beyond National Bans: The Need for ASEAN Action

For Southeast Asia, this moment reveals both potential and pitfalls. On the one hand, it highlights a niche leadership role for the region: developing practical, context-specific norms around AI harms related to gender, children and disinformation, rather than waiting for broad frameworks patterned on Western models. This is a realistic goal, especially since the region already has its own [Declaration](#) on the protection of children from online exploitation and abuse.

If ASEAN can build on this momentum, there is an opportunity to establish regional principles on AI-generated sexual and gender-based harms, deepfakes, and child protection. This could signal to AI companies that compliance with safety expectations is non-negotiable.

On the other hand, the Grok bans also highlight the risks of fragmentation. If governments unilaterally block or authorise high-risk AI tools without common standards, global companies will operate within a fragmented regulatory landscape, and vulnerable groups might remain unprotected where safeguards are weakest. Furthermore, unilateral bans could push harmful content into less-regulated spaces or drive users to circumvent restrictions, thereby undermining regulatory objectives.

To translate this moment into sustained leadership, Southeast Asian countries will need to deepen regulatory capacity and work towards ASEAN-level cooperation on enforcement. Using Indonesia, Malaysia and the Philippines' Grok decisions as case studies, policymakers can establish clear expectations for AI providers, especially for high-risk systems. This includes risk assessments, safety-by-design requirements for image tools, swift takedown obligations, and meaningful engagement with regulators before market entry.

In doing so, Southeast Asia could emerge not only as a reactive regulator of AI harms but as a contributor to global AI governance norms that incorporate regional social, legal, and political contexts.

## The Way Forward

By acting decisively against sexual deepfakes, Indonesia, Malaysia and the Philippines states have shown that meaningful AI regulation does not need to wait for comprehensive frameworks, but can proceed through targeted, enforcement-oriented responses to clearly identifiable risks. Southeast Asian states have precedent for such interventions, including handling AI-generated deepfakes during [election periods](#).

Whether this is a one-off intervention or the basis for longer-term leadership will depend on what follows. Without regional coordination and sustained investment in a regulatory capacity, unilateral bans risk fragmentation and inconsistent protection.

Conversely, if ASEAN members utilise this moment to establish shared expectations for AI providers and minimum safeguards against high-risk applications, Southeast Asia could help shape emerging global norms on AI safety and platform

accountability. In this sense, the Grok case marks not an endpoint, but a test of whether the region can turn early action into coordinated AI governance.

---

*Karryl Kim Sagun Trajano is a Research Fellow for Future Issues and Technology (FIT) at the S. Rajaratnam School of International Studies (RSIS), Nanyang Technological University (NTU), Singapore. She specialises in strategic and policy research on emerging and frontier technologies. This commentary was originally published in The Interpreter (Lowy Institute) on 19 January 2026. This adapted version is republished with permission.*

---

**S. Rajaratnam School of International Studies, NTU Singapore**  
Block S4, Level B3, 50 Nanyang Avenue, Singapore 639798

*Please share this publication with your friends. They can subscribe to RSIS publications by scanning the QR Code below.*

