Grok controversy a case study in product liability

THE decision by the Malaysian Communications and Multimedia Commission (MCMC) to block access to the AI chatbot Grok is a decisive, albeit reactive, measure. This action, taken to prevent content that creates liability under Malaysian laws, including Section 233 of the Communications and Multimedia Act 1998, serves as a necessary firebreak against the unchecked proliferation of non-consensual, sexually explicit deepfakes.

But it also underscores the timeliness of the Online Safety Act 2025 (ONSA), which came into force on Jan 1. ONSA fundamentally reshapes the liability landscape by designating social media platforms as Licensed Service Providers.

It explicitly classifies child sexual abuse material and financial fraud as "priority harmful content", which must be blocked as swiftly as possible.

While the ban addresses the immediate symptom, we must recognise that the threat is no longer theoretical or confined to foreign platforms. It is local, and it is already in our classrooms.

The case in Johor Baru early last year, where a teenager allegedly used AI to create explicit deepfake images of his schoolmates, was an early warning.

Last month, three students from a school in Muar were expelled for similar conduct.

These incidents demonstrate that the technology is accessible, easy to use and can be weaponised by anyone.

They also highlight the limitations of reactive bans. Even if we block commercial platforms like Grok, open-source models remain accessible to the tech-savvy.

Therefore, for the legal and business fraternity, the Grok controversy is a case study in product liability.

The developers of Grok deployed a tool with known vulnerabilities – specifically, the capability to "digitally undress" subjects, including minors – without adequate safeguards.

From a legal standpoint, relying on after-the-fact reporting for foreseeable harms is no longer an acceptable defence.

We are witnessing the collision between the Silicon Valley ethos of "move fast and break things" and the sovereign duty of nations to protect human dignity.

Critics often argue that strict regulation will stifle innovation and deter foreign direct investment (FDI). But this is a false dichotomy.

High-value institutional investors and serious technology majors do not seek a regulatory "Wild West"; they seek regulatory certainty.

An ecosystem where AI tools can be weaponised to generate pornography or harass citizens is inherently unstable and fraught with legal risk.

By enforcing clear standards, Malaysia is not repelling investment; it is filtering out high-risk actors and creating a safe harbour for responsible AI development.

Thus, we must pivot from reactive bans to a proactive "safety by design" framework. Any AI entity seeking market access in Malaysia should be compelled to demonstrate that safety guardrails are intrinsic to the code, not an afterthought.

Just as we require safety certifications for imported vehicles or pharmaceuticals, we must require Algorithmic Impact Assessments for generative AI tools. If a platform cannot technically guarantee that it will not generate child sexual abuse material (CSAM) upon a simple prompt, it is not "market-ready".

Our legal response moving forward must be two-pronged.

First, on the supply side, we must enforce corporate accountability. Tech giants can no longer claim neutrality; if their product design facilitates abuse, they must share the liability.

Second, on the demand side, we need urgent digital legal literacy. The public, especially the youth, must understand that using AI to generate non-consensual explicit imagery is not a "prank" or a technological experiment. It is a potential criminal offence with severe consequences under our Penal Code and the Sexual Offences Against Children Act.

The Grok ban is a necessary firebreak, but it is not a permanent solution. The future belongs to AI, but sustainable innovation requires a social license to operate.

THULASY SUPPIAH
Kuala Lumpur