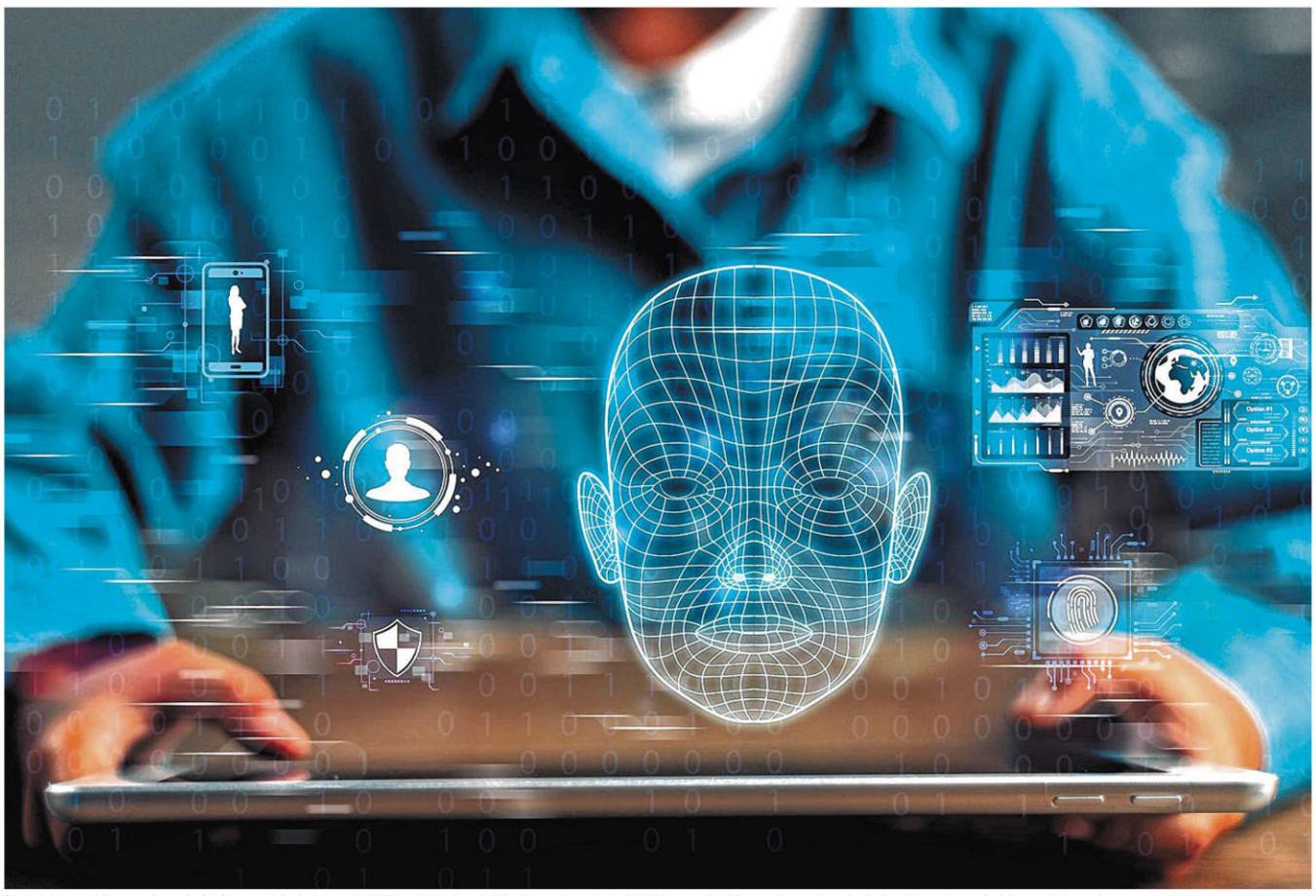
Source: *The Straits Times,* pB3
Date: 20 January 2025

Deepfake nude apps are ruining lives and have no place in app stores



Deepfake-enabling tools and platforms, with their accessibility and potential for harm, have escalated into a significant threat, particularly to women and girls, says the writer. PHOTO: ISTOCKPHOTO

The technology has unleashed a wave of harassment and exploitation. A coordinated approach is needed to tackle this escalating crisis.

Chew Han Ei

Soon, creating and sharing sexually explicit "deepfakes" will become a criminal offence in Britain. The move, announced on Jan 7 by the government, is aimed at tackling a surge in the proliferation of such images, mainly targeting women and girls. The problem isn't confined to the UK.

With artificial intelligence or AI, deepfake technology has been used to create hyper-realistic, yet entirely fabricated images and videos. Pornographic images can be digitally doctored into the likeness of an unsuspecting, innocent person and have become a widespread tool for exploitation.

Deepfake-enabling tools and platforms, with their accessibility and potential for harm, have escalated into a significant threat, particularly to women and girls, necessitating stringent regulation to curb their misuse.

It hit home in Singapore in November 2024 with the news that police were investigating deepfake nude photos of Singapore Sports School students that were created and spread by other classmates. The students' mobile devices were temporarily confiscated for forensic examination to identify and remove any remaining inappropriate content.

The incident is a stark reminder of how AI enables harassment and violation.

Stronger and more targeted legislation is certainly essential to address this issue effectively. At the end of 2024, Singapore's Ministry of Law and Ministry of Digital Development and Information launched a public consultation on proposed legislation to enhance online safety in Singapore.

Legislation is but one approach to addressing the problem, however. There are other measures to consider, all with their pros and cons.

THE TOOLS ARE AT THE HEART OF THE ISSUE

One crucial step is to regulate or ban technologies that facilitate online abuse. Mobile applications such as "undressing apps" and "nude generator apps" designed to create non-consensual sexually explicit content, have no legitimate purpose and must be removed from app stores.

In April 2024, Apple removed several generative AI apps from its app store after investigations revealed they were being used to create non-consensual nude images.

Beyond outright bans, tech companies can impose limits on searches to prevent such harmful tools from appearing prominently in search results or app stores. Measures could include deleting such apps from search algorithms, flagging them with warnings, or demoting their ranking to reduce their reach and visibility.

These actions may not fully eliminate the problem since errant developers can disseminate such apps by other means, but the measures do set a necessary precedent for accountability within the tech industry. Platforms and app developers must take responsibility for ensuring their tools are not weaponised to harm others.

They must comply with the rules to ensure user safety and transparency. Investment in better detection systems to quickly find and remove fake content is critical too.

In the global context, measures such as the European Union's online safety and AI models, and proactive reporting and rigorous AI evaluations in Australia and the UK, set strong precedents for accountability.

These actions prioritise victim protection, hold the culprits responsible and help block harmful content at its source. Enforcement must include clear penalties for platforms and developers that enable harm, ensuring accountability across the tech supply chain.

WHY OUTRIGHT BANS ARE NOT ENOUGH

However, for Singapore, amid its moves to reform enforcement and compliance measures, outright bans on harmful tools like nudification apps will not be enough. These bans often fail to address the global nature of platforms or the rapid evolution of AI technologies.

Like Pandora's box, these nudification tools have already been unleashed into the digital world and cannot simply be put back.

Developers circumvent bans by using unregulated platforms or anonymisation tools, highlighting the need for complementary strategies that combine regulation, education and technical safeguards.

The global nature of digital platforms adds another layer of complexity. Many platforms hosting harmful apps operate outside Singapore's jurisdiction, limiting the reach of local laws. Moreover, bans alone do little to address the root causes of exploitation or the demand for such tools. Without proactive education and awareness campaigns, harmful behaviours will persist, fuelled by ignorance or malicious intent.

Immediate remedies are essential for alleviating the psychological burden on victims as part of a larger framework that prevents further harm, addresses enablers and holds perpetrators to account. These measures must prioritise victims by ensuring timely removal of harmful content and reducing the burden of proving it is false.

Technical challenges further undermine bans. For instance, detection tools often fail to keep up with AI advancements, making real-time removal of harmful content a persistent challenge for regulators. These limitations highlight the need for a more comprehensive approach to combating deepfake exploitation.

FIGHTING STRATEGICALLY AGAINST DEEPFAKES

Despite these limitations, protecting minors, especially girls, from exploitation must be the priority. There is no justification for the exploitation of children.

In October 2024, a new government agency was announced to help victims quickly stop online harms without depending on the usual court-based process.

Proposed reforms include introducing a statutory complaints mechanism, administered by the new agency, to provide timely assistance to victims of online harms.

Such reforms are critical in ensuring laws explicitly criminalise the creation and dissemination of AI-generated explicit content involving minors.

But enforcement alone won't be an effective solution.

Education and rehabilitation are pivotal. For instance, a Digital Yellow Ribbon initiative that includes education on ethical technology use and pathways for reforming youth offenders would be a good start.

This offshoot of the Yellow Ribbon Project – to encourage the acceptance of former offenders in Singapore – can help address the behavioural roots of deepfake misuse and foster a culture of responsibility and empathy in digital spaces.

Singapore must also foster regional collaboration to combat the global nature of deepfake exploitation. Asean-led initiatives could establish unified standards for monitoring, reporting and removing harmful tools, amplifying Singapore's impact and influence within the region.

ADDRESSING REAL-WORLD IMPACT OF ONLINE HARM

The impact of deepfake exploitation extends far beyond the digital sphere. Victims often suffer lasting emotional, financial and reputational damage. Data from SHECARES@SCWO, Singapore's first support centre for online harms, underscores the severe toll of such abuse.

Many victims report anxiety, post-traumatic stress disorder and even suicidal ideation as a result of their experiences. The challenges are not limited to emotional scars – victims often face significant financial and psychological burdens as they attempt to remove harmful content from the internet.

Similar stories resonate globally. In the UK, 64 per cent of sexual deepfakes involve celebrities, 15 per cent depict someone the viewer knows, and 6 per cent involve the victims themselves, compounding their trauma as they face stigma and must prove their innocence by reliving the violations.

Addressing these harms requires not only supporting victims, but also ensuring that enablers of harm are held accountable through robust compliance measures. Immediate remedies are essential for alleviating the psychological burden on victims as part of a larger framework that prevents further harm, addresses enablers and holds perpetrators to account. These measures must prioritise victims by ensuring timely removal of harmful content and reducing the burden of proving it is false.

As Ms Julie Inman Grant, Australia's eSafety Commissioner, aptly said: "We cannot simply regulate our way out of these harms."

The fight against deepfake exploitation requires a coordinated approach that balances regulation, education and enforcement. By learning from global successes and tailoring reforms to local contexts, Singapore can create a safer, more ethical digital landscape.

• Dr Chew Han Ei is adjunct senior research fellow at the Institute of Policy Studies, National University of Singapore, and a board member of SG Her Empowerment.

Source: The Straits Times © SPH Media Limited. Permission required for reproduction.