Editor's note: Artificial intelligence is revolutionizing our daily lives and the ways we forge connections, transcending the barriers of time and space. The third article in this AI series examines the misuse of generative AI for nonconsensual deepfake pornography, highlighting legal gaps and the urgent need for reform.

It's my face, but it's not me

Deepfake technology has caused immense distress and concern to innocent victims and society, with experts and social groups demanding regulations to stop the menace in its tracks and prevent it from becoming a larger trend. **Wu Kunling** reports from Hong Kong.

It might all begin with a simple, everyday photo a woman shares of herself online, unaware that doing so would create a nightmare for her.

The posting ignites a barrage of indecent images of her — generated in bulk without her knowledge, saved and silently cataloged on a stranger's drive, waiting to be scrutinized by malicious eyes. The face in the pictures was hers, but the image wasn't.

In July, a male law student at The University of Hong Kong was accused of using a generative artificial intelligence tool to create indecent images of nearly 30 females by grafting their faces — mostly from social media screenshots — onto nude photographs.

Some of the victims, who were the suspect's classmates, revealed online that his computer contained about 700 images, including source photos and generated pornography. Those targeted included his university peers, primary school classmates, former secondary school teachers — some were his close friends and others were people he had met only once.

The victims said legal consultations point to a gap in Hong Kong's existing laws. While nonconsensual sharing of AI-generated intimate imagery is barred, merely creating such content without distributing it isn't a criminal offense, leaving no legal recourse to hold the student criminally liable.

The incident became one of Hong Kong's most prominent face-swapping abuse cases, highlighting how such misuse has evolved into fertile ground for more serious crimes. Yet, the regulatory framework is still wanting.

Another case in the Chinese mainland that also drew much attention involved actress Wen Zhengrong. For over half a year, she was cloned with AI from old footage and used for commercial livestreaming.

Blowing her top online, the actress hit back in comment session — "If you're Wen Zhengrong, then who am I?" — only got instantly blocked. Despite her high public profile, she admits it's hard "to prove that I am me".

Beyond China, face-swapping, notably AI-generated pornography, has emerged as one of the most alarming examples of AI technology misuse worldwide, with no nation seemingly able to emerge unscathed.

A recent survey in the United Kingdom found that 35 percent of 1,700 people aged 16 and above in England and Wales admitted to having viewed sexual and/or intimate deepfake content featuring someone they knew (14 percent) or strangers (21 percent). Conducted by police-commissioned consultancy Crest Advisory, the study also found that one in four respondents felt there's nothing wrong with creating or sharing such content, or were neutral about it.

In August, 48 state attorneys general in the United States urged technology platforms to curb the spread of nonconsensual AI-generated intimate imagery, for example, by restricting searches on "how to make deepfake pornography" and blocking terms like "undress apps" or "nudify apps".

Citing a 2023 report by cybersecurity firm Home Security Heroes, the attorneys general noted that 98 percent of computer-generated fake videos online are deepfake pornography.

"The creation of deepfakes is a grave harm to human decency and dignity, especially among our young people," says Massachusetts Attorney General Andrea Joy Campbell.

Emerging form of sexual violence

"Those photos are fake, but too real," says Doris Chong Tsz-wai, a registered social worker and executive director of RainLily — a Hong Kong-based support group combating sexual violence.

Founded in 2000, RainLily has so far helped more than 300 female sexual violence victims annually, and handled over 2,000 hotline calls with 24-hour one-stop support services, including psychological, medical and legal aid. If commissioned by victims, the organization would reach out to online platforms or coordinators of community groups, seeking the removal of nonconsensual intimate images.

In recent years, RainLily has been swamped with pleas for help from victims targeted by deepfake pornography — seven cases in the 2022-23 service year, and eight and 11 in the following two years. Although the figures were modest, Chong is worried that such acts, with basic disregard for consent and respect, could turn into a troubling trend.

In most of the cases, the victims' faces shown in publicly-shared photos were synthesized with a human body, often with exposed private parts and from unknown sources. The victims tend to be young. According to Chong, nearly all who had sought help from RainLily were in their 20s or 30s. It's also learned from social workers and teachers that nonconsensual photo manipulation occurs in secondary schools although the images produced aren't necessarily indecent.

Compared with survivors of real intimate photos leaking, deepfake victims are often "stunned" when they stumble on those images because "they never took them", says Chong.

For most of the victims seeking help, their top priority is not to lodge a report with the police, but to stop the spread of the images immediately. Although they can seek a court order to delete the images on the grounds of voyeurism, many are either unaware of this option or are too desperate to wait through the legal process.

Chong explains that many deepfake victims don't even have a clue who created the images, or when and how they might have begun circulating. This uncertainty deepens their fear, she says, noting that one victim even avoided all social media out of anxiety.

Given these distinct challenges, RainLily has had to adjust its response to such cases. Chong says this type of sexual violence has presented her team with new difficulties.

Existing laws in Hong Kong are not yet well-suited to addressing such situations — a gap that was exposed in the case involving the HKU student. Chong says she believes legislation shouldn't focus merely on whether indecent content was published or even created or not, but on the need to secure consent before using others' image, regardless of the purpose. She notes that many countries have begun enacting or are considering similar laws.

The more pressing challenge, however, is shifting public perception, Chong says.

As these incidents are relatively new, many people have yet to view them as sexual violence. This occurs not only among the public, but also in some professional assessment processes, which significantly affects how cases are handled. She notes that some victims who turned to social workers or teachers for help were greeted with comments like "it was just a joke".

"Due to such incomplete understanding, victims are told the photos are fake and not to take them seriously, while their terror over the ongoing image circulation and the growing misconceptions about them is completely overlooked," says Chong.

She urged the authorities to closely monitor areas that are not adequately covered by existing laws, update legislation in a timely manner and publicize victims' legal options. Furthermore, public education should be stepped up, especially compulsory sex education, to clarify that creating or sharing AI-generated intimate images constitutes sexual violence.

For victims, Chong advises them to seek professional support if they worry about approaching platforms or individuals alone. "Honor your feelings and don't deny yourself," she would tell them.

Governing the entire chain

The China Internet Network Information Center recently issued a report on generative AI application development in 2025, identifying voice and face-swapping as a top governance priority, ranking ahead of copyright infringement, academic misuse and tech ethics.

Liang Zheng, a professor at Tsinghua University and vice-dean of its Institute for AI International Governance, says tackling the issue is "important, but still difficult".

He notes that, from the start, major countries like China and the US have recognized the negative impact of such technology misuse, initially due to harms like disinformation that threaten public interests. But, as the technology has spread, its misuse has shifted from a societal concern to a personal one, damaging reputations and commercial interests directly.

Nevertheless, Liang says the vast, scattered user base and low technical barriers make it difficult for victims to thwart such abuse or promptly trace its origins and dissemination path. Therefore, a more viable governing approach lies in advancing both technological design and legal frameworks to establish a holistic governance model covering the entire chain — from creation to dissemination.

Within this chain, users must obtain others' personal information in compliance with China's Personal Information Protection Law and social norms, while tech companies, at every stage of the chain, must have both the obligation and capacity to exercise oversight. In this regard, he highlights two technical solutions currently being promoted.

In March, several ministries and departments jointly issued measures on labeling AI-generated content, encouraging service providers to embed invisible markers like digital watermarks in these materials. Unlike conventional watermarks, the digital ones are invisible to the human eye, but machine-readable, as they are inserted in the materials' source code. They help platforms to detect such synthetic content during subsequent upload and dissemination and, as a small consolation for victims, can also provide tangible evidence for potential judicial proceedings.

Industry and academic researchers are also refining a trigger mechanism to automatically identify sensitive material in AI-generated images and videos and take action. Liang notes that similar mechanisms for text are already well established, enabling platforms to block or restrict content containing certain keywords related to pornography or violence. He expects the technology to be effectively adapted for visual media in the near future.

Liang, however, warns that deploying these technologies is still difficult and costly, but expects stakeholders to push forward. "Building security checks earlier into the design phase may limit some product or platform functions, but it prevents many problems that could later spiral out of control," he says.

Legally, Liang says the core of addressing such infringements lies in clearly defining the responsibilities of each party in the chain, with judgments based on their specific actions, motives and the harm caused. For example, when a celebrity's face is swapped into a commercial livestream, the creator, who initiates and profits from the act, should bear primary liability. In other cases, platforms or distributors may also be held chiefly accountable.

"Existing laws provide reference principles, but judges must consider the full context of each case," says Liang. He notes that although China's laws on the rights, duties and obligations for various parties in this chain are more detailed than those in many other countries, legal frameworks will always lag behind technological development as innovation constantly produces unforeseen scenarios.

But still, regulating such misuse doesn't have to rely solely on specialized internet or AI laws. It can and should be supported by civil law provisions protecting portrait, reputation and property rights, forming a comprehensive regulatory system.

As AI-generated text, images and videos flood the internet, the old saying "seeing is believing" is losing its hold, says Liang, pointing out that false, incomplete or misleading information has always existed in the television or internet age.

What's new now, he notes, is that this era places higher demands on us — to use technologies as our tools more responsibly and continually strengthen our general knowledge and digital literacy.

"Technology keeps changing, but the principle remains the same — act responsibly and stay skeptical."

Contact the writer at amberwu@chinadailyhk.com

> "Building security checks earlier into the design phase may limit some product or platform functions, but it prevents many problems that could later spiral out of control."
>
> **Liang Zheng,** a professor at Tsinghua University and vice-dean of its Institute for AI International Governance