https://www.wsj.com/tech/cybersecurity/online-child-safety-laws-privacy-concerns-c009da8c

TECHNOLOGY | CYBERSECURITY

Stopping Child Porn Online Is a Worthy Goal. But Beware the Proposed Cure

Regulations aimed at rooting out child sexual-abuse material online could undermine many people's security and privacy

By Stuart Madnick Follow

March 20, 2025 11:00 am ET

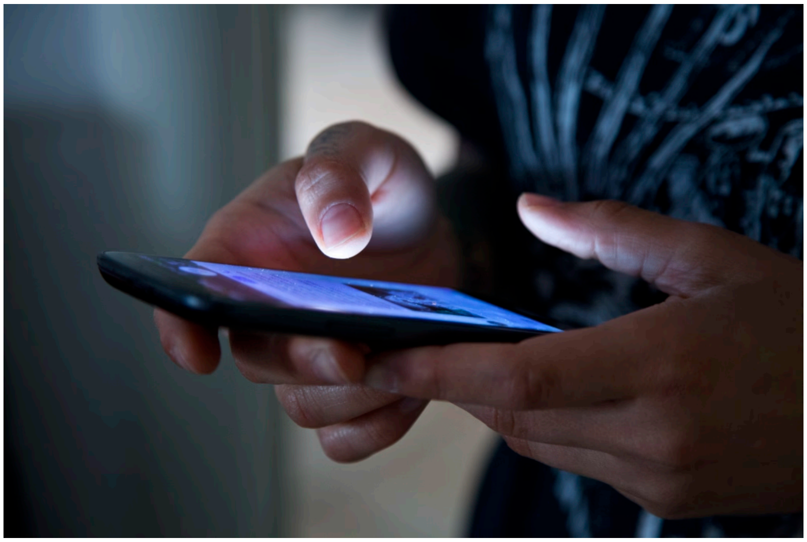

PHOTO: GETTY IMAGES

Stuart Madnick is the John Norris Maguire Professor of Information Technologies at the MIT Sloan School of Management and the founding director of the Cybersecurity at MIT Sloan (CAMS) research consortium.

The need to shut down child pornography online is one of the few things on which almost everyone can agree. So, governments around the world have proposed or enacted regulations requiring technology providers to seek out and report any child sexual-abuse material (CSAM) being shared on their platforms.

These efforts clearly are well-intentioned. But what many people may not realize is that some of the proposals could undermine the security and privacy of millions of citizens in ways that were never intended.

Here is a look at some of the potential unintended consequences, according to research I conducted with Daniel Gozman, an associate professor at the University of Sydney Business School in Australia:

Privacy breaches

End-to-end encryption uses cryptography to ensure that only you and the person you are communicating with can access the content you send. Even the digital platform you are using can't see it. Encryption helps prevent cybercriminals from being able to use your information even if they steal it.

The Experts are a group of industry and academic thought leaders who weigh in on topics covered in the The Journal Report.

But if technology providers are going to be held liable for illegal photos shared on their platforms, they will need a way to bypass encryption to scan users' content. If service providers change their systems to provide such capabilities—say, by creating a master key or a "backdoor" point of entry—the risk grows that personal or sensitive information could fall into the hands of unauthorized third parties.

At a minimum, suspect photos selected by the scanning system—whether ultimately determined to be illegal or not—must be reviewed by people, possibly many people, who weren't authorized by the owners and don't need a search warrant. That represents an immediate breach of privacy and increases the risk that personal or sensitive material could fall into the wrong hands and be exploited.

False accusations

Automated detection of illegal photos isn't perfect. A Facebook study of 150 accounts that were reported to authorities for having alleged child sexual-abuse material found that 75% of the accounts were incorrectly flagged. Another study of a method often used to detect CSAM—called perceptual image hashing—found its accuracy rate could be as high as 95%.But on WhatsApp alone, where as many as 4.5 billion images are shared daily, a 95% accuracy rate could result in more than 200 million images a day being incorrectly identified. Innocent individuals could be subjected to police investigation and other consequences.

Government overreach

Although the stated goal of these regulations is to search for child porn, once governments have the ability to decrypt their citizens' communications and information, the risk grows that they could use it for less noble purposes—say, to search out critics of the government. Politicians would have unprecedented levels of access to track any user's personal life and activities without giving users any choice or requirement for a warranty.

Vulnerability to cybercrime

Although intended to be used only by the "good guys," the decryption tools used by digital platforms could be stolen and exploited by "bad guys" to gain access to otherwise protected information. It has happened time and again, including in a recent breach, when cyberattackers exploited backdoors that telecommunication companies created to lawfully share information with authorities.

The concerns above might be viewed as acceptable if it was guaranteed that all child sexual-abuse material would be eradicated. But not only do the current recognition techniques often mistakenly flag legal content, they often miss illegal content, too. That's because there are clever techniques that perpetrators can use to disguise such content. One study showed that slight modifications resulted in 99.9% of images bypassing commonly used scanning algorithms.

Although the goal of rooting out child sexual-abuse material online is undeniably worthy, policymakers need to carefully weigh the risks of large-scale scanning of private content. There may be better options. Rather than creating ways to override the privacy
protections of end-to-end encryption, it might be better to build on and strengthen the
existing laws and regulations and by educating the public to be more alert.

Write to Stuart Madnick at reports@wsj.com.