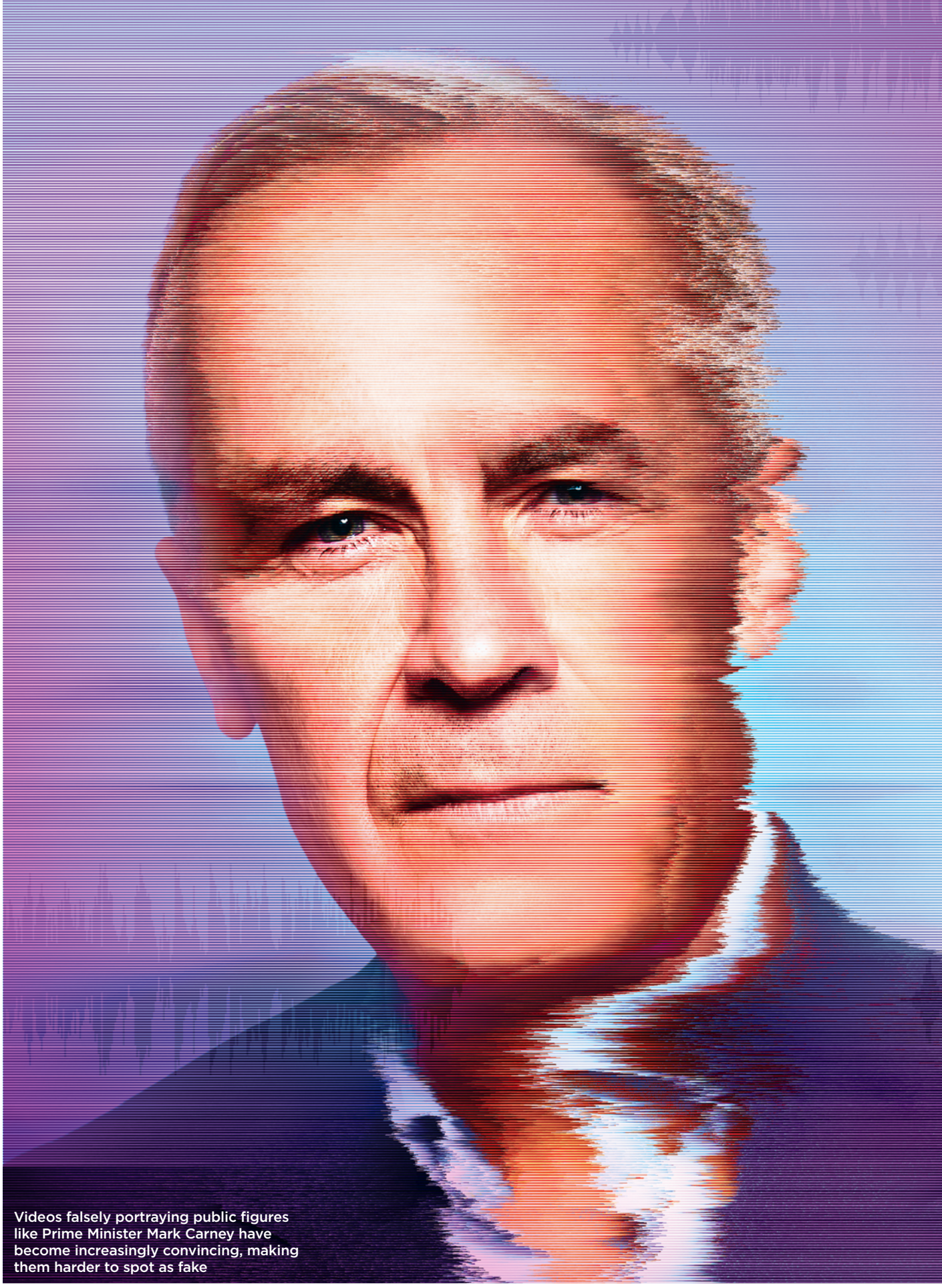


Videos falsely portraying public figures like Prime Minister Mark Carney have become increasingly convincing, making them harder to spot as fake

PHOTOGRAPH BY **RÉMI THÉRIAULT**

Look Again

The rapid spread of deepfakes has unnerved businesses and cybersecurity experts alike. As these AI-generated videos become more lifelike, CPAs will need to stand guard against the deception. **By John Lorinc**

Earlier this year, Instagram users with an interest in investing found a highly compelling post in their feeds: Abby Joseph Cohen, a legendary Goldman Sachs investment strategist who now teaches at Columbia Business School, popped up in a video, promoting three deeply undervalued stocks that, she seemed to be predicting, were set to experience enormous gains.

However, as a subsequent investigation by *Fortune Magazine* revealed and Goldman Sachs confirmed, the videos were "deepfakes," constructed by generative AI software using existing online video and audio clips of Joseph Cohen. *Fortune* found numerous examples of deep-faked investment execs starring in other social media pump-and-dump schemes. "AI-enhanced scams pose significant risks to investors," observes Bonnie Lysyk, former Ontario auditor general and currently the Ontario Securities Commission's (OSC) executive vice-president for enforcement. "The use of AI can divert those investors away from legitimate sites to fraudulent sites," she adds, citing research carried out by the OSC and other securities regulators.

Although deepfake technology is most closely associated with phony celebrity videos and AI-generated porn, there's a growing list of examples of the technology being used by international criminal organizations intent on defrauding corporations as well as gullible investors. Bad actors now equipped with inexpensive and easily accessed software tools have found novel ways to use deepfakes to mess with internal financial controls, hiring and onboarding systems, authentication protocols thought to be secure and a range of other systems.

Indeed, such uses of generative AI are proliferating rapidly, and now pose a daunting challenge to corporate cybersecurity teams and their advisers, who have spent years training managers, employees and executives to be aware of phishing schemes and ransomware attacks. Almost overnight, the list has grown to include the criminal use of deep-faked voice clips—a.k.a. spoofing—to gain access to consumer banking services. "Emerging technologies such as AI have a rapid development and deployment strategy," says Peter Hargitai, PwC Canada's national digital risk solutions leader. "Therefore, a disciplined and thoughtful approach is needed to timely embed cyber and data trust protocols and controls."

Trust under siege

Many of us have probably tried one, with mixed results: those news site quizzes that test your ability to distinguish a real image from an AI one. There are the most obvious clues—too many fingers, unnaturally smooth skin, missing shadows, etc.—but generative AI's ability to learn from its mistakes has meant that the clues have become ever more difficult to spot without sophisticated scanning tools.

Not surprisingly, a 2025 report by the Association of Certified Anti-Money Laundering Specialists ranked criminal use of AI as fourth of the ten most worrisome financial crime threats currently in circulation. "The use of generative AI for malicious purposes continues to profoundly reshape the threat environment," the authors wrote. "The sheer speed, scale and sophistication of criminal exploitation—notably with malware, social engineering and video/voice cloning, sometimes referred to as 'deepfakes'—make this one of the most significant threats to financial crime functions."

For CPAs providing guidance on risk, cybersecurity and data governance, the deepfake threats facing corporate clients are not just formidably fluid but potentially expensive. A frequently cited and bracing

Abby Joseph Cohen, former chief investment strategist at Goldman Sachs, was targeted in a deepfake stock scam ▼

Employees must know how to respond swiftly to AI-enabled deception before financial or reputational damage occurs

public example involves a Hong Kong–based finance worker with Arup, the U.K.-based engineering giant, who was summoned to a virtual meeting with other members of staff and told to transfer US$25 million to an external account. He followed the instructions. As it later transpired, the entire virtual meeting session had been convincingly manufactured by deepfake fraud artists. (Arup declined to comment.) "The incident speaks to how advanced the capabilities of these tools have become and continue to evolve," says Melissa Robertson, CPA Canada's principal for research and thought leadership. "It's getting much harder to be able to differentiate when it is an AI versus when it is a person."

"Generative AI tools are much more accessible to create convincing fraud," adds Alan Mak, BDO's national forensics practice leader. "The earlier phishing emails, in hindsight, were pretty easy to detect. You get an email that on the header says it came from our CEO and click onto it. It's pretty obvious that the actual email address is not the address; it's something else. Now we're talking about either audio or visual fakes, or both, that would be harder to detect because it looks like you're speaking to, or it feels like you're speaking to, the person you think you are. But in reality, you're not."

While incidents such as the one in Hong Kong get a lot of attention, voice "spoofing" is another pressing concern because so many companies, like banks and telcos, have adopted voice authentication to make it easier for customers to access their accounts remotely. This year, OpenAI announced its latest platform could create precise voice replicas based on just 15 seconds of authentic audio, a capability that effectively weaponizes the seemingly limitless amount of audio now available online, through social media, podcasting, YouTube and more. (Deepfake voice prints have been implicated in some notorious consumer frauds, such as the grandparent scam, where fraudsters place calls to elderly relatives asking them to immediately transfer large sums to deal with an emergency situation involving a beloved grandchild.)

Voice spoofing has rapidly upended what many cybersecurity experts regarded, until recently, as a robust verification tool. Customers could record themselves, and the algorithm was trained to recognize their voices when they called in to do their banking. That's not the case anymore. Earlier this year, *The Wall Street Journal* reported that Chase Bank's voice authentication system was "fooled" during a test, although the bank insisted it has other backup verification systems. According to a 2024 report by cybersecurity solutions provider BioCatch, over nine in ten banks are now thinking about eliminating the feature in favour of the most robust dual-factor authentication systems—a statistic in wide circulation. "If you reach out to the contact centre of some of your current service providers," says PwC's Hargitai, "and you've signed up for their voice authentication, you can practically go all the way through and start transacting. Appropriately protecting and safeguarding telephony technologies is a must."

As concerning are the use of deepfakes in cybersecurity fraud aimed at senior finance executives and repositories of valuable corporate data. A recent Deloitte survey found that over a quarter of respondents had been targeted by fraudsters looking to steal financial data within the last year. In other cases, deepfake technology is being used to create false online customer IDs using stolen, personally identifiable information (PII), as well as AI-generated photos imported to deep-faked documents, such as driver's licences and passports.

Recruitment processes, in turn, have become a particularly vulnerable portal because so much hiring and onboarding is now done virtually. "During COVID, we saw people pretend to be someone else to get a job, but they're really not that person and they're not even qualified," says Marilyn Abate, a partner in KPMG's Toronto forensics team. "With deepfakes, they can make it look like [someone else]. And if they're working remotely all the time, you'll never know who the real person is." The payoff for the fraudsters, she says, is access to the company's

PHOTOGRAPH BY GETTY

internal systems. "Sometimes the criminals will plant people in organizations to commit insider fraud."

One U.S. cybersecurity service provider recently posted a story about how, when seeking to fill a routine IT position, it inadvertently hired a North Korean programmer looking to infiltrate the organization; the individual was using a stolen U.S. ID that had been enhanced with AI, but eventually left a trail of increasingly suspicious digital bread crumbs. In other cases, fraud artists posing as executive recruiters for well-known firms will lure job seekers through networks like LinkedIn, then demand applicants pay a fee to continue the process. As Abate adds, "I've seen examples of larger organizations that's happened to—that [fraudsters] infiltrated the organization through this technique and were able to commit fraud, steal money and do other bad things."

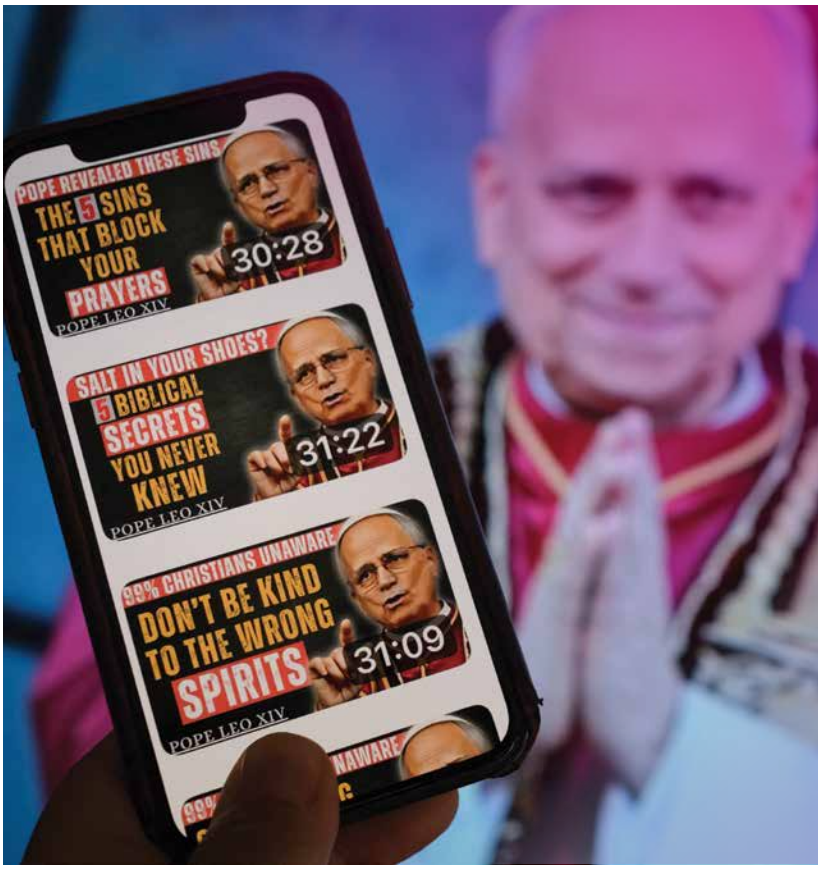
International responses

Given a technology as dynamic as generative AI and a globalized threat environment, regulators and enforcement agencies have found themselves playing catch-up, as usual, but there's evidence that most are trying to respond in real time.

In late 2024, the U.S. Treasury Department's Financial Crimes Enforcement Network (FinCEN) released a detailed alert to advise financial institutions on how to detect the types of deepfake threats that should be included in suspicious activity reports collected under banking legislation.

The document not only detailed recent developments in generative AI but itemized numerous reported techniques used by criminals to create customer accounts with fake documents made via deepfake technology or the use of deepfake audio/video in phishing emails. FinCEN's guidance included a list of nine red flags and related clues, such as inconsistency between customer documents, the use of third-party webcams during video authentication calls and, of course, suspicious patterns of account activity. Among its list of best practices: live verification checks initiated by the financial institution in which the customer is prompted to confirm their identity using video or audio.

Here, the Canadian Securities Administrators' network of provincial regulators, as well as the Office of the Superintendent of Financial Institutions (OSFI) and the Canadian Centre for Cyber Security, have all issued similar warnings to market participants, issuers and investors. In a recently published alert, OSFI and the Financial Consumer Agency of Canada also listed the types of threats posed by generative AI, including incursions aimed at smaller banks and credit unions, which traditionally avoided attacks because their size makes them less attractive to criminal gangs.



▲ AI-generated videos of Pope Leo XIV are spreading online fast, racking up views as platforms struggle to rein them in

"Between the securities commissions and the other regulators in Canada and the investment players, I think there's work being done to try and address the problems that we're seeing with impersonation," says Lysyk. Yet Hargitai argues that rules governing corporate disclosure to capital markets for AI-related cybersecurity risk mitigation and related governance activity are still evolving in Canada, and more is needed to address the evolving risk vectors.

The enforcement end of the story is even less encouraging. Last year, the European Union, which takes a proactive approach to data protection, passed the AI Act. It includes provisions requiring developers and deployers to reveal to end users when they're interacting with AI (for example, via chatbots or deepfakes), according to an extensive summary by University of British Columbia law professor Benjamin Perrin. He also points to the first Canadian court case involving deepfakes of child sexual exploitation, based on charges brought against a 61-year-old man who not only collected online child porn, but pleaded guilty to creating at least seven videos with so-called deepfake technology.

PHOTOGRAPH BY GETTY

"The rise of increasingly genuine-looking audio and visual deepfakes, now powered by AI, has added one more tool to the arsenal of threat actors"

As for deepfakes used in fraud or cyberattacks on companies, Canada doesn't have much in the way of significant legal protections, says Lisa R. Lifshitz, a Torkin Manes partner who chairs the firm's technology, privacy and data management practice. "We don't actually have bespoke deepfake legislation," she says, describing the unnerving experience of watching presenters at tech law conferences produce highly convincing deepfakes in minutes. "We have certain federal laws, such as the Criminal Code, that prohibit fraud, identify theft and harassment, for example, and copyright laws that offer certain protections, but we don't have deepfake legislation," says Lifshitz.

She noted that existing Canadian privacy laws offer limited recourse if the deepfakes do not reveal personal information. Certain provinces have recognized the statutory tort of unauthorized use of someone's likeness and the tort of "appropriation of personality." The tort of "false light" has been upheld in Ontario, but Lifshitz adds that financial penalties are so low that they provide little protection.

Another potential cause of action is defamation. The Quebec Civil Code confirms that individuals have a right to the respect of their reputations and privacy, and a deepfake that appropriates or uses a person's name, image, likeness or voice for a purpose other than the legitimate information of the public would be a violation under the code.

Lifshitz and other experts note that problems or frauds associated with deepfakes must become a routine part of what corporate cybersecurity teams monitor, in terms of identification, technology defences, data governance, training and other forms of risk mitigation. "Unfortunately, the rise of increasingly genuine-looking audio and visual deepfakes, now powered by AI technologies, has added one more tool to the arsenal of threat actors that require constant vigilance," she says. The wrinkle, adds Mak of BDO, is that the technology "enables bad actors to bypass some of our traditional controls."

In some cases, the prevention and risk-mitigation tools are actually embedded in classic approaches to financial governance. The notorious case of the Arup manager in Hong Kong revealed a weakness in the company's internal systems, says Robertson of CPA Canada. "One individual should not be able to go start to finish to transfer $25 million." The incident, she adds, "emphasizes the need to have proper internal controls in place so that you can prevent situations like that from happening."

In some cases, says Abate of KPMG, executives and managers have begun to use an additional level of authentication, such as a "safe word" that can be used when someone posing as a company official has directed another employee to make a large transfer. "It's something that only you and I know so somebody else can't mimic it," she says.

The turbocharged sophistication of deepfakes has also made regular cybersecurity training that much more important. To some extent, Mak acknowledges, it's always going to be a matter of playing catch-up. "Our clients are running fraud training and phishing training two, three, four times a year, which means there could be anywhere from three months' to twelve months' delay in employees or staff hearing about what's happening, whereas the bad guys are constantly evolving." Best practice involves follow-up tests and consequences for those who repeatedly fail.

"Having the right level of investment in this area," adds Hargitai, "is going to be an ongoing battle." The tool kit includes security-by-design approaches to IT systems, such as minimizing the number of people with access to PII or customer accounts. Another risk-mitigation tool involves technologies that can screen for voice spoofing or add digital watermarks to authentic video, flag deepfakes on platforms like Zoom or check the source of a video feed to see if it traces to a known address. Yet an evaluation of AI uses in capital markets, published

this year by the International Organization of Securities Commissions, warned that these techniques are "not comprehensive and can be evaded."

Opinion on the efficacy of such tools remains mixed. "They can be effective," says Mak. "My caution is that you should not rely on them 100 per cent for obvious reasons. They provide a good first line of defence, but my advice is generally to use multiple lines of defence." Lifshitz is more skeptical: "Using digital watermarks may be helpful now, but in the longer term, AI technology will make them obsolete."

Because the technical sophistication means deepfakes are much harder to detect, Abate stresses that employees in large organizations shouldn't necessarily be trying to decode a suspect image themselves; all they need to do is be alert to something that doesn't seem quite right. "That's why you need the training because they can't just rely on the system," she says. "You need people's eyes. You don't have to train [employees and managers] to know what to do with [a suspect deepfake]. Just raise your hand and tell someone."

The last move involves incident response tailored to the deepfake threat in particular, as one cybersecurity expert told a U.S. Securities and Exchange Commission investor forum this year. "The rise of AI-driven scams means businesses must be prepared for deepfake-enabled fraud, synthetic identity attacks and AI-enhanced phishing schemes," said Perry Carpenter, a deepfake researcher. "Organizations should update incident response plans to include rapid fraud assessment procedures, forensic AI analysis capabilities and clear internal escalation paths. Employees, finance teams and security personnel must know how to… respond swiftly to AI-enabled deception before financial or reputational damage occurs."

Future certainties

Two things are clear in this volatile domain: one, that deepfakes and their uses will become inexorably more realistic, persuasive, numerous and therefore difficult to identify; and two, that the rise of this brand of cyber threat will produce lucrative new business lines for cybersecurity consultants and other advisers. Deloitte's Center for Financial Services estimated last year that GenAI could enable fraud losses to reach US$40 billion in the United States by 2027, so firms need to invest in prevention, training and detection or they'll face bracing losses associated with the mischief that deepfakes can inflict. "Unfortunately," Mak muses, "people don't really take it seriously until they've been hit. But once they've been hit, everyone takes it very seriously." ◆

DIGITAL IMPOSTORS

How deepfakes can slip into financial services

By John Lorinc

Last year, the Financial Services Information Sharing and Analysis Center, a Virginia-based cybersecurity not-for-profit, published a list of scenarios of existing deepfake scams that pose a risk to the financial industries. These include:

C-suite impersonations targeted at investors, employees or consumers in the service of pump-and-dump schemes, fraud and access to personally identifiable information

Scams that leverage voice print identification to bypass banks' customer verification controls, gaining access to consumers' accounts to defraud them

Deepfake videos that replicate financial advisers or investment bankers' likenesses, using their identities and reputations to commit fraud or other forms of online misrepresentation

Deepfakes that impersonate third-party advisers who have relationships with financial service firms, which can then be used to gain access to or exfiltrate money from these firms

PHOTOGRAPH BY ISTOCK