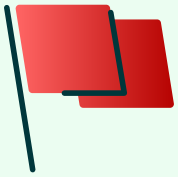
billtrust

WHITE PAPER

The Dark Side of AI

How finance leaders can combat deepfakes and phishing schemes



For every glowing adjective lavished on Generative AI, there seems to be an equal number of harrowing warnings.

As finance teams explore the use of AI to streamline processes and remove repetitive tasks from their to-do lists, criminals are using the same exact tools to commit check fraud, create far more polished phishing emails, poison data, and even extort money through imposter scams.

Take the most memorable caveat surrounding this impossible-to-ignore technology: beware of deepfakes.

While in the past, AI-generated deepfakes were so crude that they would fool only the gullible, today's tech savvy fraudsters can spoof famous individuals and convince even sophisticated viewers that they're "seeing" a celebrity, a politician or a familiar business leader.

The dark side of GenAI encompasses everything from harmless antics to financially devastating imposter scams. To guard against falling victim to generative AI fraud, companies are investing in ensuring their computer systems—and their workforces—are not compromised by malicious AI.

Generative AI's biggest threats

$2.7B

The estimated amount of losses due to imposter scams in 2023

Source: The Federal Trade Commission

In 2023, consumers lost more than $10 billion to fraud, the highest amount associated with fraud ever recorded, according to a Federal Trade Commission (FTC) report.[1]

One of the prime culprits in the rise of fraud? "Digital tools," according to Samuel Levine, Director of the FTC's Bureau of Consumer Protection.

Take, as an example, imposter scams, many of which occur because criminals have learned to use artificial intelligence to spoof the identities of real people. The FTC estimates that in 2023, impostor scams accounted for $2.7 billion in losses.

Another swiftly growing type of generative AI-dependent fraud is BEC, or business email compromise. According to the FBI, there were 21,489 BEC complaints in 2023, with adjusted losses exceeding $2.9 billion. BEC scams often involve falsified invoices, typically arriving through a business email account that has been compromised.

BEC scams often involve falsified invoices, typically arriving through a business email account that has been compromised.

In other BEC scams, a bad actor might request W-2 information from an unwitting victim or from an HR or payroll professional. Because the requests have been carefully constructed to look legitimate, employees too often reveal personal data.

Although these scams can be perpetrated without artificial intelligence, ChatGPT and other generative AI systems are making the requests appear far more professional than they otherwise might have—and they are therefore less likely to be flagged as fraudulent.

Gone are the days when an email with typos or bad grammar was a tip-off that it wasn't coming from a credible source. Now, even non-native speakers can write an impressive sounding email, with no grammatical or spelling errors.

Top frauds

1. Imposters
2. Online shopping and negative reviews
3. Prizes, sweepstakes, and lotteries
4. Investments
5. Business and job opportunities

Source:
Federal Trade Commission[2]

GenAI brings check fraud to a new, alarming level

Check fraud is an old crime that's gotten a dramatic makeover with generative AI.

Criminals are using tried-and-true methods of perpetrating check fraud, including stealing checks from mailboxes and washing checks to alter payee and dollar totals. What's different in 2024 is that bad actors are "confirming" their stolen identities by AI-enabled means. A recent media report shows that one site—OnlyFake—is allegedly using generative AI to create as many as 20,000 falsified ID documents each day.[7]

One way to tackle check fraud is by taking fuller advantage of new payment methods, says Lisa Devashrayee-Oaks, Billtrust's Lead Product Marketing Manager. "We try to help our customers get their customers away from paying with checks and instead use ACH or systems that are more secure," she says.

AI-generated profile photos from thispersondoesnotexist.com

How to identify deepfakes

What would you do if your UK-based CFO invited you to a video call and then asked you to pay more than $25 million to one particular individual? If you behaved like a Hong Kong worker earlier this year did, you'd make the payment. This scam succeeded because the staff members on the call were deepfake recreations, according to the Hong Kong police.[3,4]

Often, deepfakes are perpetrated over Zoom, Microsoft Teams, or another videoconferencing platform. The technology has evolved to the point where the person a viewer is "seeing" may be a fraudster who doesn't even resemble the individual appearing on screen.

Bank of America also warns against shallowfakes (sometimes known as cheapfakes), which tend to use crude methods to deceive. Here, think filters or airbrushing.[5]

To guard against everything from deepfakes to cheapfakes, business leaders need to recognize when something is amiss. Here are some telltale signs:

1. **Out-of-sync actions.** The gestures or lip movements of a person speaking on camera may not match the audio.
2. **Pauses and long periods without blinking.** In both audio and video, longer-than-usual pauses may indicate that an original has been tampered with.
3. **Patchy skin tone.** Bank of America notes that deepfakes are often detectable by examining an individual's jawline or ears.[6] If facial features seem blurred or the skin tones don't match, that's a red flag.

1,077,501

The number of phishing attacks in the 4th quarter of 2023

Source: APWG

Phishing schemes are on the rise

Rather than going through the enormous work of breaching firewalls, figuring out passwords, thwarting antivirus software, then hacking a system, criminals often prefer to talk an employee into supplying banking or other key data through social engineering, making entry far easier. The endgame of social engineering is often phishing attacks that deceive individuals into revealing personal information or doing something ill-advised like installing malware.

The numbers of phishing incidents are growing fast. In the fourth quarter of 2023, 1,077,501 phishing attacks were observed, according to the Anti-Phishing Working Group, or APWG. APWG notes that attacks against social media platforms exploded in late 2023 (accounting for 42.8% of all phishing attacks that took place).

Another alarming trend is "spear phishing." While many phishing attempts are scattershot and may be addressed to numerous individuals, spear phishing attempts succeed at higher rates because they use data specific to an individual to appear more convincing.

Emerging AI threats

Generative AI is breathing new life into age-old types of fraud, as it's making some entirely new threats possible. In fact, 49% of security leaders surveyed in 2023 worry about criminals using AI or machine learning (ML) to bypass security precautions.[8]

Here are a few specific generative AI threats identified by the new ATLAS (Adversarial Threat Landscape for AI Systems) framework:

- **Data poisoning.** Say a bad actor intentionally adds false information into a machine learning model. That actor could then "poison" detection algorithms, allowing fake transactions to be authorized.
- **Evasion attacks.** If hackers understand how your facial or voice recognition systems work, they can fool these systems by, say, using a photo instead of an actual face for ID purposes.
- **Backdoor AI.** If a fraudster gains access to where AI models are stored, the criminal could upload a lookalike program designed for a completely different purpose. One worst-case scenario would be installing a new model that doesn't detect certain threats, resulting in a security breach.

How finance leaders can identify GenAI threats

If criminals are using cutting-edge technology tools to deceive individuals, then finance leaders and AR and AP professionals must use data more wisely, too.

"We provide near real-time analytics, so you can see if anything looks 'off' in your data," says Devashrayee-Oaks. Monitoring patterns in data is one excellent way to detect fraud early.

Here are a few other ways to identify when generative AI is possibly being used for malicious ends:

- **Educate your workforce about deepfakes.** AI experts believe that regular training sessions and workshops can aid employees in telling the difference between genuine and manipulated content.
- **Restrict use of generative AI at work.** A 2024 Cisco study shows that 27% of companies have banned, at least temporarily, the use of generative AI at work because of privacy and data security risks.[9]

Hysteria? A gross overreaction? Not necessarily. The study found that 48% of participants admitted to entering private company information into GenAI tools.

- **Stress test your systems**. Just as AI can be taught to write a grammatically perfect email, these systems can be trained to identify malicious prompts.
- **Engage in simulations.** Phishing simulation exercises are one way to teach employees what a social engineering attempt may look like. A lighthearted way to demonstrate how easy it is to be deceived by deepfakes? Have employees test themselves on a site like this.

The future: New threats emerging

Experts often describe an "arms race" between those attempting to use generative AI for good and those using it for malicious ends.

Because generative AI is so new, the learning curve about its criminal uses can be steep. That said, it's critical that finance leaders take the time to educate themselves—and others. Increasingly, finance leaders are telling customers that they'll never ask for a password or even communicate about certain subjects by text. Letting your customers know how you do and do not communicate empowers them to identify when a call or text seems suspicious.

In the brave new world of generative AI, forewarned is forearmed.

> Rest assured that Billtrust is staying on top of emerging threats so finance leaders know what they're up against. We partner with finance teams to help them take full advantage of the promise of generative AI without falling prey to emerging scams and fraud attempts.

To-dos for warding off Generative AI scams

Take advantage of **modern payment methods,** such as ACH systems

Get familiar with your data throughout the O2C cycle, then **monitor those data patterns**

Enlist an AI expert to train your staff. Often, a trusted payments solutions partner can keep you abreast of the latest fraud tactics—and how to combat them.

References

1. Federal Trade Commission. "As Nationwide Fraud Losses Top $10 Billion in 2023, FTC Steps Up Efforts to Protect the Public." https://www.ftc.gov/news-events/news/press-releases/2024/02/nationwide-fraud-losses-top-10-billion-2023-ftc-steps-efforts-protect-public
2. Federal Trade Commission. "As Nationwide Fraud Losses Top $10 Billion in 2023, FTC Steps Up Efforts to Protect the Public." https://www.ftc.gov/news-events/news/press-releases/2024/02/nationwide-fraud-losses-top-10-billion-2023-ftc-steps-efforts-protect-public
3. CNN. "Finance worker pays out $25 million after video call with deepfake 'chief financial officer." https://www.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html
4. OECD.AI Policy Observatory. "Deepfakes and AI: How a 200 Million Scam Highlights the Importance of Cybersecurity Vigilance." https://oecd.ai/en/incidents/69834
5. Bank of America. "How to protect your business from deepfakes." https://business.bofa.com/en-us/content/cyber-security-journal/deepfakes-business-risks.html
6. Bank of America. "How to protect your business from deepfakes." https://business.bofa.com/en-us/content/cyber-security-journal/deepfakes-business-risks.html
7. 404 Media. "Inside the Underground Site Where 'Neural Networks' Churn Out Fake IDs." https://www.404media.co/inside-the-underground-site-where-ai-neural-networks-churns-out-fake-ids-onlyfake/?_hsmi=292891802&_hsenc=p2ANqtz-9N4Ern8WGWMlY2vK46j5Rw-MOai28yFe5M1c7WYsvWCNdmOvX3_s4kXKCJBt9S3UtaiR_RC6QHmfEzdbpPw1j0WAjtbEA
8. Security Magazine. "Software supply chain compromise was fourth most frequent attack." https://www.securitymagazine.com/articles/99641-software-supply-chain-compromise-was-fourth-most-frequent-attack
9. Cisco. "More than 1 in 4 Organizations Banned Use of GenAI Over Privacy and Data Security Risks." https://newsroom.cisco.com/c/r/newsroom/en/us/a/y2024/m01/organizations-ban-use-of-generative-ai-over-data-privacy-security-cisco-study.html

billtrust

Learn more

Visit billtrust.com or contact our sales team.

ABOUT BILLTRUST

Finance leaders turn to Billtrust to get paid faster while controlling costs, accelerating cash flow and maximizing customer satisfaction. As a B2B order-to-cash software and digital payments market leader, we help the world's leading brands move finance forward with AI-powered solutions to transition from expensive paper invoicing and check acceptance to efficient electronic billing and payments. With more than $1 trillion invoice dollars processed, Billtrust delivers business value through deep industry expertise and a culture relentlessly focused on delivering meaningful customer outcomes.

CORPORATE HEADQUARTERS

11D South Gold Drive
Hamilton Township, New Jersey 08691
United States

SACRAMENTO

2400 Port Street
West Sacramento, California 95691
United States

GHENT

Moutstraat 64 bus 501
9000 Ghent
Belgium

AMSTERDAM

H.J.E. Wenckebachweg 200-III
AS 1096 Amsterdam
Netherlands

KRAKÓW

ul. prof Michała Życzkowskiego 19
3 piętro
Kraków 31-864
Poland

PM-BT-WP23-v1