Insights, News & Events

AI IS MAKING PHISHING SCAMS ALMOST IMPOSSIBLE TO SPOT – 5 STEPS EMPLOYERS SHOULD TAKE TO COMBAT LATEST THREAT

Insights
Sep 2, 2026

AI is Making Phishing Scams Almost Impossible to Spot – 5 Steps Employers Should Take to Combat Latest Threat

Artificial intelligence tools have made it dramatically easier for scammers to produce convincing phishing messages, with the speed of change significantly outpacing the defenses most employers currently have in place. AI-generated messages are now fluent and well-written, eliminating the grammatical errors and awkward phrasing that used to serve as tell-tale warning signs. At the same time, cybercriminals are increasingly automating their entire attack plans rather than just the messages themselves, using AI-powered tools to create personalized lures and execute attacks at a scale and speed that was not possible even a year ago. All this means that your existing phishing training and verification practices may be out of date, even if you updated them as recently as last year. Here's what's happening and five steps you should consider taking.

Warning Signs Have Changed

For years, employers trained workers to watch for the tell-tale signs of a scam: awkward phrasing, grammatical mistakes, language that didn't sound like it came from a native speaker or a real colleague, and return email addresses that looked like gibberish. That advice is no longer reliable. Security awareness training built around spotting flaws needs to become building awareness of situations. Employees are now facing scam attempts that are

Related People

Daniel Pepper, CIPP/US
Partner

303.218.3661

Michael Steele
Chief Information Officer

404.240.5858

far more sophisticated than what most training programs were built to address.

- Generative AI tools can now produce **fluent, natural-sounding messages** in virtually any language, eliminating the errors that used to give scammers away.
- Scammers are increasingly training AI tools on real corporate marketing and communication content, allowing fraudulent messages to **closely mirror the tone and style** of a company's actual internal or customer communications.
- AI tools can quickly **research a target company and its employees**, building a detailed picture of reporting lines, job responsibilities, and internal terminology that can be used to make a fraudulent message feel credible.
- Cybercriminals can also generate **look-a-like email domains** designed to fool the eye.

"Device-Code Phishing" is the Fastest Growing Trend

Beyond more convincing writing, cybercriminals are increasingly automating entire attack sequences from research through execution. One emerging technique, known as **device-code phishing**, exploits a legitimate authentication process originally designed for devices that cannot easily accept a typed password.

In these attacks, the attacker starts a sign-in of their own and sends the employee a short code. The employee is sent to the real sign-in page, not a counterfeit one, and enters the code. When the employee finishes signing in and clears multifactor authentication, the identity provider issues the access token to the attacker's waiting session. Nothing is spoofed and nothing is intercepted.

The employee has authorized the attacker's device which is why multifactor authentication does not stop this attack: the requirement was satisfied, not bypassed. The resulting tokens can keep an intruder in the mailbox for weeks unless someone revokes them, and a password reset alone does not do that.

This particular attack method has surged this year. Cybersecurity firm Huntress reported a nearly 15-fold increase in device-code phishing attempts during the first four months of 2026 compared to the second half of 2025.

Service Focus

AI, Data, and Analytics

Data Protection and Cybersecurity

Privacy and Cyber

Resource Hubs

AI Governance Hub

Much of that growth has been tied to subscription-based phishing kits that bundle identity-theft tools, prebuilt phishing infrastructure, and AI-generated content into an accessible package. This allows attackers with little technical skill to launch sophisticated campaigns.

What's Next?

Researchers who study these attack patterns have projected that fully automated, AI-driven email compromise attacks could become common before the end of 2026 (read more here, subscription required). They also expect more complex and damaging attack types following as soon as 2027.

5 Steps Employers Should Consider Now

Given how quickly the underlying warning signs have changed, employers should revisit both training content and internal verification practices. Here are five steps to take:

1. Update phishing training to reflect current scams. Employees should no longer be told that spotting bad grammar or awkward phrasing is a reliable defense. Your training should instead teach them to recognize situations such as unexpected codes, urgent requests, or unusual payment requests, and to verify the source before acting, regardless of how polished the message appears.

2. Build out-of-band verification into policy. For wire transfers, credential resets, and changes to payroll or direct deposit information, require confirmation through a separate channel, using a number from your own directory and never a number supplied in the request itself. Add one instruction most policies are missing: a familiar voice on that callback is not confirmation, and the legitimacy of a message proves nothing. Voice cloning is cheap and convincing now, and video is close behind. The control is in using the channel you choose and the contact information you already have on file.

3. Extend scrutiny beyond email. Because attackers are increasingly using messaging platforms like WhatsApp, Slack, and text messages to impersonate executives or IT staff, your verification practices should apply across all communication channels, not just inboxes.

4. Educate employees on newer attack techniques. Device-code phishing and similar methods that bypass multifactor authentication are not yet widely understood outside of

security teams. Defensive tools now use AI to counter AI-driven attacks, but technology will not solve this alone. Bringing IT security and HR together to explain how these attacks work can help employees recognize unfamiliar but increasingly common request patterns.

5. Treat this as a joint IT and HR responsibility. Because these attacks are designed to exploit workplace trust and reporting relationships, effective defenses require collaboration between security teams and HR (or legal) teams. Fraudulent direct deposit changes are the most common HR-facing version of this attack, and the legal consequence surprises employers. In most states, wages sent to an account controlled by a criminal have not been paid to the employee, so the employer generally has to pay again, and some state wage payment statutes add penalties for the delay. Tighten the process before you are litigating it by requiring out-of-band verification for every banking change, holding the change for one pay cycle where your system allows, and sending a confirmation notice to the employee's address of record whenever bank details are changed.

Conclusion

Fisher Phillips will continue to monitor developments and provide updates as warranted, so make sure you are subscribed to Fisher Phillips' Insight System to get the most up-to-date information direct to your inbox. If you have questions, contact your Fisher Phillips attorney, the authors of this Insight, or any attorney on our Data Protection and Cybersecurity Team.