Communications of the ACM

Opinion

Artificial Intelligence and Machine Learning

AI Rewrites the Rules Of Phishing, Cybercrime

How artificial intelligence can supercharge attempts to penetrate your security.

By Gaurav Belani

Posted Dec 18 2025

Gaurav Belani
Social Engineering 2.0
Doi: 10.1145/3772374
https://bit.ly/46TQ2Jg

In early 2024, a Hong Kong-based clerk at multinational consulting engineering firm Arup was duped into transferring about $25 million to scammers who used AI to impersonate the company's CFO, and other senior executives, in a live video meeting. The fraud only became apparent when the employee checked in with headquarters in London.

This is social engineering 2.0.

In the old days, phishing emails were usually riddled with typos and bad grammar, or at least had subtle tells like domain names with typos. With a little cybersecurity education, you could spot them a mile away. Now, AI studies your team's LinkedIn profiles, writes flawless messages, and builds a ruse so tailored it feels like it's from your best client. They might even use deepfakes or voice clones to be even more convincing.

AI is the ultimate force multiplier for cybercriminals, because it makes scams faster, cheaper, and more convincing at scale. It can automate tasks that criminals used to do manually—or would never have invested the time to do manually. It can personalize attacks using scraped data, mimic voices, faces, and writing styles with staggering accuracy.

In this post, let's take a look at how AI is rewriting the rules of phishing and cybercrime.

How AI Is Supercharging Phishing Attacks

The next wave of phishing and AI attacks is already here, driven by technologies such as agentic research, deepfakes, and machine learning that are making threats more convincing than ever.

Here's how today's cybercrooks are using AI to create large-scale phishing campaigns that are almost impossible to spot:

1. **Hyper-Personalized Lures**

   Forget the generic "Dear Customer" email. AI can scan your social media, past press releases, and even your GitHub commits, and use that data to craft a message that feels hand-written for you. It can reference your recent vacation, your company's latest product launch, or even an inside joke from your social media feed.

2. **Deepfake Audio and Video**

   Voice cloning used to work well only when fed hours of recording for training. Now, a 30-second clip is enough to mimic your boss's tone and accent. Combine that with AI-generated video, and you get full-blown fake Zoom calls convincing enough to move millions.

3. **AI-Driven Chatbots That Never Slip**

   Old scam chats used to break character quickly; AI chatbots don't. They respond instantly, adapt to your tone, and can hold a believable "customer service" conversation while quietly harvesting your personal info.

4. **Real-Time Language Switching**

   AI translation is now good enough to scam you in flawless Spanish in the morning, Hindi by lunch, and Japanese in the evening. No awkward phrasing, no giveaway grammar mistakes.

5. **Attacks at Machine Speed**

   What once took days of research and prep can now happen in minutes. AI can spin up thousands of unique phishing campaigns at the same time, each one tailored to its victim.

Case Studies: Recent AI-Powered Scams

It used to be just a sci-fi nightmare scenario, but today, AI phishing is real, and it's costing companies millions.

We've already touched upon this one, but the Hong Kong phishing scam that targeted an employee at Arup deserves a deeper dive. The employee was tricked by deepfake versions of her CFO and colleagues into transferring HK$200 million across 15 transactions. The case has been widely reported and confirmed by the Hong Kong police.

Every face and voice was AI-generated. The employee thought she was following her

CFO's orders on a video call. The money was gone before anyone realized.

Why did the scam work so well? For starters, it wasn't a simple shady email. It was a full-on video call with recognizable faces and voices. All fake, but super realistic. The person on the screen looked and sounded exactly like the CFO.

And when someone who looks like your boss tells you to do something, most people just do it.

They played the classic pressure game too, talking about how urgent the transactions were. That shuts down the little voice in your head saying, "Hmm, should I double-check this?"

But here's the thing: There were red flags; the employee just needed to know what to look for. For one, why were there so many big payments going out to previously unknown accounts, especially ones based overseas? And why did such a major transfer come out of nowhere, with no heads-up, no prior discussion?

Even the video might've had tiny glitches. Maybe a weird blink, or slightly off lighting. Deepfakes aren't perfect (yet). And if the company had a rule that big payments need multiple approvals, well, that rule got skipped. That's a big fat clue something's wrong.

Similarly, back in 2019, criminals used AI-based voice cloning to impersonate the CEO of a German company. An executive at a subsidiary, a U.K.-based energy firm, thought he recognized the CEO's tone and accent and transferred about $243,000 to a fake supplier before realizing it was a scam.

The fake CEO asked for a quick payment to a supplier. The cloned voice perfectly mimicked the CEO's accent, tone, and cadence, making it sound authentic over the phone. It also matched a real deal the company was working on, so it didn't feel strange at the time.

Furthermore, the caller sounded urgent, saying it had to be done right away to close the deal. So the executive did what was asked of him.

But again, there were warning signs hiding in plain sight. The payment was going to a new account the company hadn't used before. That alone should've raised a flag. And if the exec had just pinged the real CEO through a known channel like email or internal chat, the whole thing might've unraveled.

Also, that mix of pressure and secrecy is a classic social engineering tell. It's meant to convince the victim to override logic and act quickly, circumventing protocols. And while the voice clone was highly convincing, these tools still mess up now and then. Weird pauses, flat emotion, audio-video sync issues, lack of movement, and tiny slip-ups are clues if you're paying attention.

It's evident that AI is helping to create new attack surfaces. AI scams succeed because they blend real context with hyper-realistic impersonation. The more familiar they feel,

the harder they are to doubt, which is exactly why double-checking through a trusted, separate channel is non-negotiable.

Detection and Defense: What Works in the AI Era

To defend against these dark arts, there's a need to fight fire with fire.

To begin with, you can't rely on human eyes alone. Modern security tools use machine learning to spot anomalies in tone, sender patterns, or user behavior. They can catch subtle signs, like if a coworker emails you at 3:00 a.m. or if the tone feels off. Things humans might miss, machines can flag.

Next, the old "trust but verify" approach is dead. Verification comes first, every time. The new "zero trust" policy means every device, user, and request gets double-checked whether they're inside or outside your network.

There's also tech that tracks your typing rhythms, how you move your mouse, and even how you talk. So if the "CEO" types like a different person, the system can tell. That's behavioral biometrics for you.

But tech alone isn't enough. If social engineering is now 2.0, user awareness should also be 2.0. Teams need exposure to AI-generated phishing simulations so they learn to spot scams that look perfect. Drills should cover video calls, chat platforms, and phone scams —not just email.

Also, use out-of-band verification. Big request? Large sums of money? New account details? Always confirm through a separate, trusted channel like a known phone number or in-person meeting. This one extra step shoots most scams in the head.

Finally, if someone smells something phishy, they should know exactly what to do. Prepare an incident playbook with clear steps that empower your team to lock it down fast and limit the damage.

Stronger Weapons

AI has changed phishing forever. Attacks are now faster, more convincing, and nearly impossible to detect with the old "look for typos" playbook.

Deepfakes, voice clones, and AI-driven chatbots have expanded the scammer's toolkit, making even seasoned employees vulnerable. Defending against these threats means matching AI's speed and sophistication with tools like anomaly detection, zero trust verification, and realistic phishing simulations that go beyond email.

Above all, verification is your strongest weapon. Confirming big or unusual requests through a trusted, separate channel can stop most scams in their tracks. Audit your phishing defenses now, before AI tests them for you.

About the Authors

Gaurav Belani *is a Senior SEO and Content Marketing Analyst at Growfusely, where he specializes in crafting data-driven content strategies for technology-focused brands.*

Join the Discussion (0)