

What is AI-assisted phishing and social engineering?

AI-assisted phishing and social engineering use generative models to make deception faster, cheaper, and more convincing: fluent and localized phishing emails, tailored lures built from public data, synthetic personas, and cloned voices or video for phone-based (vishing) and video-call fraud. The techniques are not new, but AI removes the usual tells like bad grammar and generic wording, and it scales personalization. The defenses are the fundamentals done well: verify unexpected requests on a trusted channel, use phishing-resistant multi-factor authentication, and control any action that moves money or access.

HOW IT WORKS

01 How attackers use AI

The practical uses are concrete:

- Fluent, localized lures: grammatically perfect emails in the target's language, matching a brand's tone.
- Personalization at scale: tailor each message using public or leaked data about the target.
- Synthetic personas: build believable profiles and histories for long-cons.
- Voice cloning for vishing: clone a voice from a short sample to make a convincing phone call.
- Deepfake video: impersonate an executive on a video call to authorise a payment or access.

Each of these removes a red flag people were taught to spot.

02 Why it works

The old guidance is losing its edge. Telling people to watch for bad grammar fails when the grammar is perfect. Both the volume and the quality of attacks rise at once. Voice and video cloning defeat the instinct that recognising someone's voice or face proves who they are. And the skill needed to run a convincing campaign drops, so more attackers can do it.

HOW TO DEFEND

- Verify unexpected requests on a channel you already trust, by calling back a known number, not the one in the message.
- Require phishing-resistant multi-factor authentication such as passkeys or FIDO2, so a stolen password is not enough.
- Enforce out-of-band and dual control for anything that moves money or changes access.
- Publish SPF, DKIM, and DMARC so mail that forges your domain fails.
- Train people on the new reality: verify identity through a trusted channel, do not rely on an accent or a face.

SOURCES

- [1] Adversarial Threat Landscape for AI Systems
- [2] Cybersecurity Best Practices
- [3] AI Risk Management Framework

Find out whether your team still falls for the lure.

securelayer7.net/learn/ai-security/ai-assisted-phishing

[Open online](https://securelayer7.net/learn/ai-security/ai-assisted-phishing)