CYBERSECURITY 101 · eBuilder Security

AI in Cybersecurity

AI is now on both sides of every attack. Here is what changed, what it costs and the one habit that stops the most expensive version of it.

WHAT IT IS

AI in cybersecurity is the use of machine learning and generative models by defenders and attackers alike. Defenders use it to spot unusual behaviour, sort alerts and respond faster. Attackers use the same technology to write fluent phishing in any language, clone a voice from seconds of audio, find flaws in code and automate parts of an intrusion. It rarely invents a new attack. It makes the familiar ones cheaper, faster and far harder to dismiss as fake.

WHY IT MATTERS

More than 80%

Of observed social engineering activity worldwide was AI-supported by early 2025.

ENISA Threat Landscape, 2025

73%

How often listeners correctly identified deepfake speech, with training barely helping.

PLOS ONE, 2023

94%

Of 804 surveyed leaders expect AI to be the biggest driver of change in cybersecurity this year.

World Economic Forum, January 2026

WHAT HAPPENED

In January 2024 a finance employee at the Hong Kong office of engineering firm Arup joined a video call with the chief financial officer and colleagues he recognised. Every one of them was AI-generated. He made 15 transfers to five bank accounts totalling HK$200 million, roughly 25 million US dollars. No internal system was compromised. The whole attack ran on a video call and a plausible story. (Hong Kong Police, February 2024, confirmed by Arup, May 2024)

RED FLAGS TO WATCH FOR

Spelling mistakes and stiff grammar are gone, so watch the behaviour instead of the wording.

- Urgency together with secrecy, a deal that cannot wait and must not be discussed.
- A channel that changed, arriving on WhatsApp, a new number or a personal address.
- A process being skipped, where the payment or access grant avoids the normal route.
- Short scripted answers on a call, with the other party steering away from open conversation.
- Fluent detail about the business and nothing only a colleague would actually know.

THE ONE RULE THAT STOPS IT

Verify every unusual or high-value request through a second channel you already hold, before anything moves. Never through the channel the request arrived on.

THEN BUILD THE HABIT AROUND IT

- Call back on a number held in your own finance records, never one supplied in the request.
- Treat voice and video approval as insufficient on its own, no matter who appears on screen.
- Agree a challenge question for executives and finance staff, then practise using it.
- Name the sanctioned AI tools in writing and state what may never be pasted into them.
- Keep an inventory of AI systems, agents and suppliers, marking which can act on data.

MYTHS AND FACTS

MYTH

You can tell an AI voice if you listen carefully.

FACT

In a 2023 PLOS ONE study of 529 listeners, people identified deepfake speech correctly 73% of the time and training improved that only slightly.

MYTH

The EU delayed the AI Act, so nothing applies until 2027.

FACT

Only the high-risk obligations moved to December 2027. Article 50 transparency duties, including chatbot disclosure and deepfake labelling, apply from 2 August 2026.

MYTH

AI security is a job for the IT department.

FACT

NIS2 Article 20, transposed by Cybersäkerhetslagen, makes the management body responsible for approving and overseeing security measures, with personal accountability for board members.

THE LEGAL DUTY

Cybersäkerhetslagen (SFS 2025:1506) has applied since 15 January 2026, with an early warning due within 24 hours and a full notification within 72 hours whether or not AI was involved. From 2 August 2026 the EU AI Act also requires you to disclose AI interaction, mark AI-generated content and label deepfakes.

GET HELP

eBuilder Security helps Swedish organisations build the verification habits, monitoring and training this guide describes. Read the full guide at the link below.

Read the Full Article →