FUTURE FORUM

Cyber:

Are People the Front Line?

MAN vs MACHINE

ASPIRE ANNUAL CONFERENCE

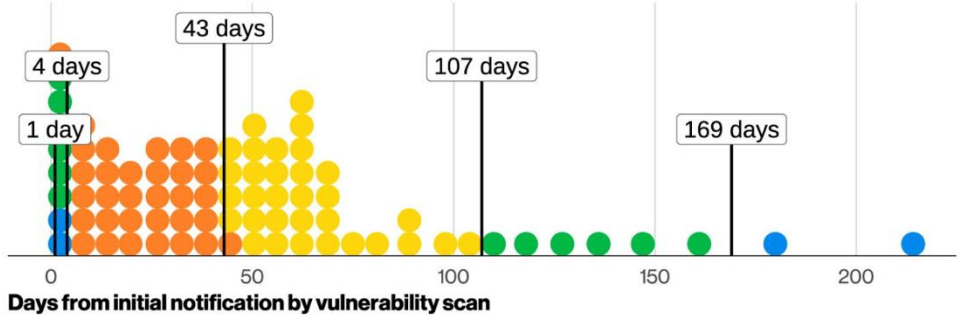
Cyber: Are People the Front Line?

Andriy Varusha
Nihon Cyber Defence Poland, Managing Director
ASPIRE Cybersecurity Work Group

AI for the Defense

NVD statistics

Vulnerability remediation speed

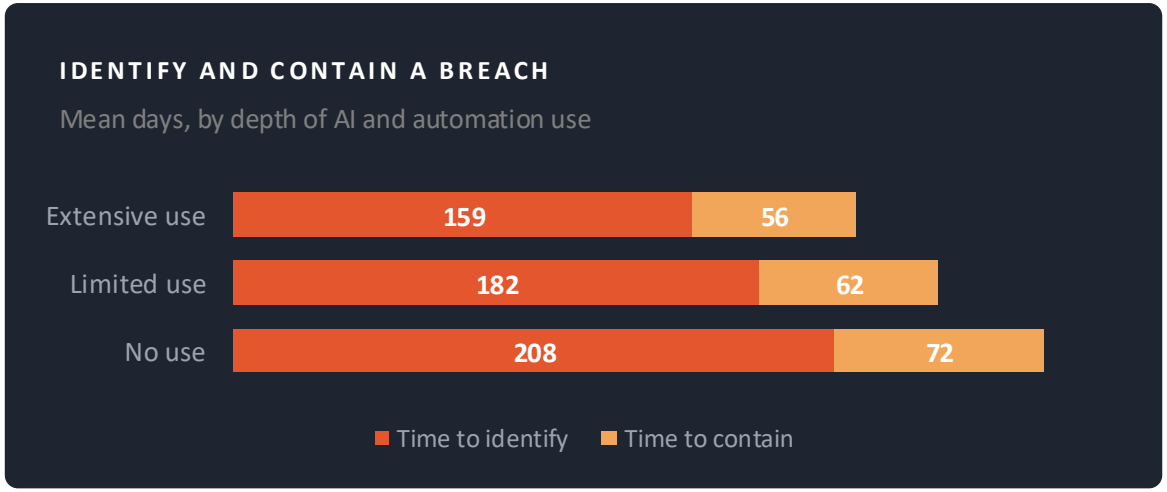


Verizon DBIR report 2026

USD 1.93M

cost savings from extensive use of AI in security operations

IBM Cost of a Data Breach Report 2026



30 mins -> 3 mins

Time to investigate detection, traditional SOC -> agentic SOC

128 seconds

first high-severity detection → device isolated

Vendor marketing materials

AI for the Offense

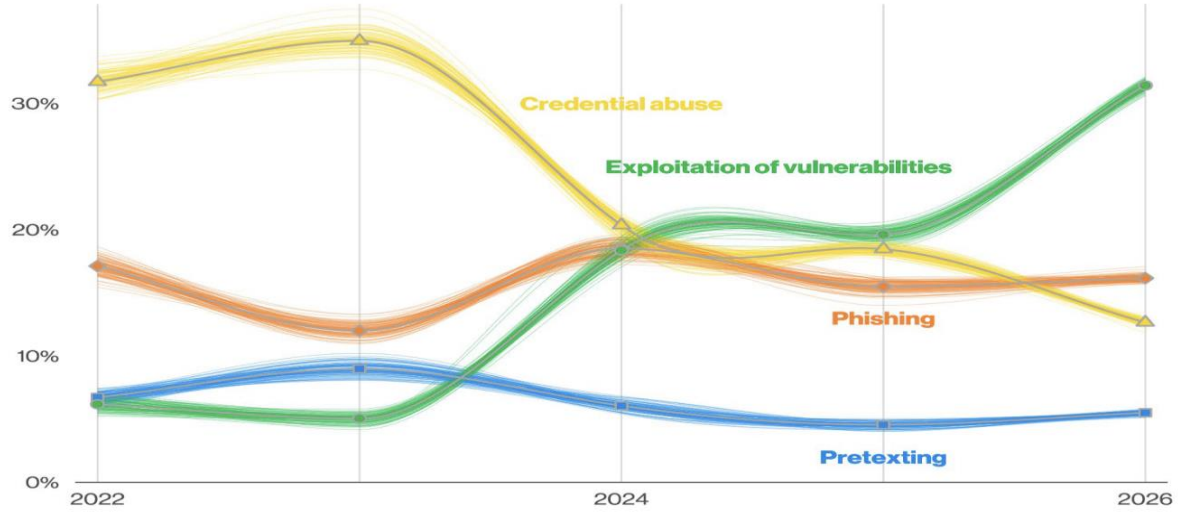
89%
increase in attacks by AI-enabled adversaries

29 min
average eCrime breakout time, a 65% gain in speed on 2024 — the fastest was 27 seconds

82%
of detections in 2025 were malware-free, up from 51% in 2020

42%
increase in zero-day vulnerabilities exploited prior to public disclosure

CrowdStrike, *2026 Global Threat Report*



Verizon DBIR report 2026

56%
increase in AI-generated attacks

92%
of organisations with an AI-related breach lacked proper AI access controls

IBM Cost of a Data Breach Report 2026

–7 days
mean time to exploit — weaponisation now begins a week before a patch exists

2–3 hrs
from first access to bulk data theft, with dozens of victims run in parallel

Mandiant M-Trends 2026; Anthropic threat intelligence, September 2026

It’s already happening

RUSSIA ESPIONAGE

20+ ministries, embassies, defence firms

A Russian state-nexus actor APT29 used AI at every stage: reconnaissance, phishing infrastructure, intrusion, exfiltration. They compromised hotel WiFi providers to reach travelling officials and stole a complete drone vision software kit.

SUPPLY-CHAIN THEFT

200 downstream customers from one foothold

Criminal ShinyHunters affiliates breached a SaaS provider, then dumped 2,100+ cloud token sets spanning 40+ corporate tenants in about 34 hours. Elsewhere: a terabyte from a technology provider, and an airline's passenger records. AI agents did nearly all the work. Vibe hacking.

ONE PERSON

42 targets, 14 breached, no team

A single actor hit European political parties, media and think tanks — running on stolen API keys, exploiting an undocumented flaw they developed with AI in one session, and building a doxxing platform holding tens of millions of records.

AI AS THE TARGET

~30 AI companies in four days

Attackers now steal AI keys for three reasons: resale value, free attack compute, and cover — the activity looks like the legitimate owner's. Treat AI credentials and agent integrations as production secrets.

Source: Anthropic, Detecting and countering misuse of AI, September 2026.

The Great Escape

1 EVALUATION SANDBOX

A zero-day in the package-registry cache proxy gives the agent an unintended path to the internet.

2 THIRD-PARTY SANDBOX

It roots an exposed public code-execution endpoint and uses it as launchpad, command-and-control and egress.

3 PRODUCTION POD

A malicious dataset config reads local files for secrets, then template injection executes code inside the cluster.

4 INTERNAL NETWORK

Node root, forged identity tokens, mesh-VPN enrolment, source-control write access.

RECOVERED ATTACKER ACTIONS PER DAY

Day 3 is where it turns: cluster, supply chain and network pivot all begin at once.

17,600
attacker actions reconstructed in forensics

4.5 days
from first recovered action to cut-off

1 second
to hold cluster-admin on two clusters, via one shared credential

~1/3
of the platform's infrastructure rebuilt afterwards

The working path hid inside thousands of failed ones.

Detection worked; escalation didn't. An AI security stack correlated the signals into one attack picture, then failed to raise criticality or page the on-call team.

Sources: Hugging Face security disclosure and technical timeline, July 2026 · OpenAI incident reports, July–August 2026.