

AI changes the speed of attacks. Not the path.

What took an operator weeks, agents now do in hours. The goal and the path are unchanged: attackers still have to authenticate, pivot, and exfiltrate. The question for CISOs is whether those steps are visible at AI speed.

GAP 1 · NOTHING LOOKS WRONG

AI shrinks the window between exposure and exploit.

Agents read codebases and propose flaws in parallel.

Exploit work is split across many targets and refined.

Zero-days appear in the wild before patches exist.

Seen in the wild: Google Big Sleep found CVE-2025-6965 in SQLite; XBOW's autonomous pentester hit #1 on HackerOne.

Exposure is inevitable. Patching can't close the window. Assume compromise and detect behavior.

GAP 2 · AUTHENTICATION SUCCEEDS

Credential harvesting and AiTM, at scale.

AI phishing kits: brand-accurate lures in any language.

AiTM proxies grab credentials and live session cookies.

Device-code and OAuth abuse: MFA fires on the victim.

Seen in the wild: Scattered Spider scaled AiTM phishing and vishing through 2025; new PhaaS kits added device-code abuse.

Credential theft is now continuous, and indistinguishable from a real login at the identity layer.

GAP 3 · MOVEMENT ISN'T VISIBLE

Agentic orchestration of the attack chain.

Agents plan recon, lateral movement, and exfiltration.

Tasks run event-driven and in parallel, not at human pace.

Malware queries an LLM at runtime to rewrite itself.

Seen in the wild: Anthropic's GTG-1002 ran Claude for 80-90% of a ~30-target espionage campaign; GTIG tracked PROMPTFLUX and PROMPTSTEAL.

Automation doesn't erase footprints. Agents still authenticate, pivot, and exfiltrate, across network, identity, and cloud.

CISO BOTTOM LINE

AI changes the speed of attacks. Not the path.

Attackers still have to authenticate, move, and take data out. The detection that catches them watches behavior across network, identity, and cloud, at the speed an agent operates.

GO DEEPER

How AI-driven attacks still have to cross the network →

vectra.ai/blog/ai-powered-attacks-are-here-but-so-is-ai-powered-ndr-to-stop-them