

Shifting to Prevention-First Security to Stop Faster AI Threats

Closing exposure before attacks start,
without adding complexity



All Rights Reserved. © 2026 Bitdefender. All trademarks, trade names, and products referenced herein are the property of their respective owners. The information contained in this document is confidential and only for the use of the intended recipient.

You may not publish or redistribute this document without advance permission from Bitdefender.

AI-Enabled Attacks: What Has Actually Changed

For several years the industry has been told that offensive AI is about to change everything. A good deal of that reporting ran ahead of the evidence, and IT leaders were right to treat it carefully.

The evidence has caught up. In a survey of 1,200 IT and cybersecurity professionals, only 18% still believe AI attacks are mostly industry hype.¹ Most now report encountering the effects directly:

- ▶ 59% say their organization has experienced social engineering they attribute to AI.
- ▶ 56% say the same thing about malware-based attacks.
- ▶ 53% say AI is helping attackers more than it is helping defenders.

The more useful finding is what did not change. Working with Anthropic's safeguards team, the Verizon 2026 Data Breach Investigations Report analysed 793 threat actors flagged for misuse of AI. The median actor used AI across 15 techniques already catalogued in MITRE ATT&CK.² Known moves, executed faster. Microsoft, Google and the Center for Strategic and International Studies reached the same conclusion independently: the real effect of generative AI is that existing methods become faster, cheaper and more accessible.³

The tradecraft did not get cleverer, the economics of running it collapsed. What AI compresses is time: the hours needed to find a weakness worth using, the gap between a vulnerability becoming known and an exploit existing for it, and the effort of chaining several small weaknesses into one working path. It also drives the cost of a failed attempt close to zero. A cheaper attacker runs many crude attempts rather than one careful campaign, and each failure is still traffic to inspect and an alert to judge on your side.

Public authorities are arriving at the same place from different directions. In July 2026 the EU Agency for Cybersecurity found that frontier models compress the vulnerability management lifecycle and the attack chain from discovery through to exploitation, raising the prospect of negative time-to-exploit, where a working exploit exists before a fix is available.⁴ Three months earlier, Singapore's Cyber Security Agency advised organizations to prepare for the same compression, reporting that frontier models can cut the time taken to identify vulnerabilities and engineer exploits from months to hours, while noting no sign of misuse so far.⁵

AI-enabled attacks are arriving faster and in greater numbers.
The techniques behind them have not fundamentally changed.

"Security fundamentals matter more than ever in the age of AI."
ENISA, ENISA's view on Cybersecurity in the Frontier AI Era, July 2026

Which raises the obvious question. Which fundamentals, and in what order?

Expanded Attack Surfaces Create the Gaps that AI Attacks Exploit

A breach is almost never one failure. It is a chain, and a typical one looks something like this:

- ▶ An internet-facing service nobody was scanning.
- ▶ An unpatched flaw on it.
- ▶ A credential carrying more rights than the job needs.
- ▶ An administrative tool sitting on a machine whose user has never opened it.
- ▶ A control that was licensed but left at its default.

Each link is individually unremarkable. Together they are a path from outside to the data that matters.

Expanded attack surfaces create opportunities for attackers to get into or move across networks, and they exacerbate the challenge of addressing AI-enabled attacks that find vulnerabilities and exposures and exploit them much faster.

Two links in a typical chain deserve particular attention, because of the exposure they introduce and because current vulnerability management and hardening do not address them properly.

Internal AI Usage

Employees and developers have adopted chatbots, coding assistants, browser extensions, local model runtimes and agents faster than anyone has catalogued them. Adoption ran ahead of governance, and security became accountable for a surface it cannot see.

Abuse of Tools You Already Trust

The Verizon dataset contains a finding that rarely makes headlines: 83% of privilege escalation incidents involved no Common Vulnerabilities and Exposures (CVE) exploitation at all.² Once inside, attackers mostly did not need a vulnerability. They used what was already on the machine.

Bitdefender Labs analysis of 700,000 security incidents found that attackers abuse legitimate tools already present in the environment in 84% of major attacks.⁶ In most cases this administrative tooling was sitting on endpoints whose users never touched it.

The Challenge of Having a Small Team

Mid-market organizations do not have the security expertise and staff to manage different complex solutions, mitigate vulnerabilities, and reduce risk.

Asked about attack surface reduction specifically, 57.0% of professionals agree they need to reduce their attack surface. The barrier they name first is not doubt about the value:

- ▶ **38.0%** point to the overhead of maintaining hardening rules and exceptions.
- ▶ **34.6%** cite resource constraints.
- ▶ **33.8%** cite visibility gaps.¹

How Are Organizations Addressing the Challenges?

The security industry is responding to these new challenges. Two responses dominate right now, and each one is a reasonable answer to a part of the problem.

Continue With Detection and Response, Accelerated by AI

Security did not always work this way. Antivirus, firewalls and endpoint protection were built to keep attacks out, and for years that was the strategy. The industry then accepted that sophisticated attacks would get past preventive controls, took on an assume-breach posture, and spent the last decade building better and faster ways to catch an incident before it became a breach. Endpoint Detection and Response (EDR), Extended Detection and Response (XDR) and Managed Detection and Response (MDR) answered the problem of the time addressing novel, more sophisticated and sometimes faster attacks.

The answer to the rising speed and volume of AI-enabled intrusion attempts has been to augment detection and response with agentic AI capabilities that enable faster and sometimes machine-speed response.

This is where the overwhelming majority of cybersecurity AI investment has gone, and if implemented correctly, with a human in the loop, this approach genuinely compresses the time between an alert and a decision. AI-enabled detection and response capabilities are part of the solution to machine-speed threats but it does not fully address the problem. If a swarm of malicious agents is easily able to compromise preventive layers and gain a foothold in the environment, reactive response can still be overwhelmed as attackers win enough time to exfiltrate data and cause damage to organizations.

Conventional Vulnerability and Attack Surface Management

Most organizations are already aware of the importance of vulnerability and patch management and have tools and processes to reduce the potential of an attacker exploiting a recent CVE that was not yet patched. These processes work to make businesses that adopted them a harder target to hit, but as the number of vulnerabilities has rapidly increased each year it has become clear that teams cannot efficiently reduce their time-to-patch and keep up with attackers' speed.

In addition to patching, many teams rely on cybersecurity hygiene activities such as hardening endpoints, and application control to reduce risk and these are based on periodic reviews and updates to identify drift and adapt policies to new threats and business needs.

To increase visibility across newer attack surfaces, teams will also add new point solutions, despite this resulting in several consoles to reconcile. IBM's Institute for Business Value found the average organization runs 83 separate security solutions from 29 vendors, and that 52% of executives name complexity as the biggest impediment to their security operations.⁸



Where These Approaches Fall Short

Each response runs into a different limit in the mid-market.

Triage Capacity Does Not Scale With Attack Volume, and Some Attacks Do Not Wait

AI on the detection side speeds up the handling of each alert. It does not reduce how many arrive.

AI-enabled detection and response capabilities are part of the solution to machine-speed threats, but it does not fully address the problem. If a swarm of malicious agents is easily able to compromise preventive layers and gain a foothold in the environment, reactive response can still be overwhelmed as attackers win enough time to exfiltrate data and cause damage to organizations.

Bitdefender Labs found close to 900 malicious skills on ClawHub, the public registry for the OpenClaw agent framework, roughly a fifth of everything published there.¹⁰ In the most widespread campaign the agent tells the user a fix is required, the user pastes the single line they are given, and the rest runs unattended: a dropper lands, the macOS quarantine flag is cleared, and a stealer sends high-value files out. Because activities look legitimate and trusted tools are involved, the risk is that detection and response will not stop the attack before exfiltration occurs.

Accelerating detection and response by integrating agentic AI in MDR services and security tools is critical, but it cannot be the first line to address AI-enabled threats.

Current Hardening and Attack Surface Reduction Cannot Keep Up

Missing or Fragmented Visibility: Organizations today have only partial visibility across their critical attack surfaces: many still focus primarily on software exploits and ignore other exposures related to shadow IT, over-privileged access to cloud accounts, or unnecessary access to risky admin tools. In some instances, visibility is fragmented across several tools, making it difficult to construct a single picture of all critical exposures.

Poor Prioritization: Ranking by CVE score answers how bad a flaw is in the abstract, not whether an attacker could reach that asset and if it is actively exploited in the wild. Conventional vulnerability and attack surface management tools cannot identify the real attack paths and chains of exposures that can be exploited by AI agents or human actors to reach sensitive assets. Given the ever-increasing number of vulnerabilities and attack surfaces and the speed of modern attacks, effective prioritization has become crucial to identify the exposures that matter and fix them before they can be exploited.

Slow mitigation: Classic attack surface reduction is a rule set. Someone writes it, exceptions accumulate, the business changes, and it drifts out of date between reviews, while patch windows lengthen and time-to-exploit trends toward zero. Getting exposures reliably fixed, every cycle, before attacks can take advantage of them, is one of the most challenging tasks to address.

The new tools bought to close each of those gaps make the same trade. Each purchase helps address a new attack surface, but this approach adds complexity that rapidly becomes unmanageable for organizations with lean IT or security teams.

The challenge facing mid-market organizations:
How to close visibility, prioritization and mitigation gaps without adding
complexity or cost

Addressing the Challenges Without Adding Complexity

The following steps outline how organizations can remove attack opportunities and reduce the strain on response capabilities as well as the risk of data breaches, all without adding complexity and tools that require additional staff to maintain them.

1. Adopt Prevention-First as Your Security Foundation

A control that closes a path in advance behaves identically whether one attacker probes it, or a thousand do. Prevention is volume-independent; triage is not. Given that AI's principal effect is more attempts arriving faster, that is where investment compounds rather than merely keeping up.

Two practices carry that load. Attack surface reduction removes the tools, privileges and paths an attacker would use, before anyone tries to use them. Exposure management finds and ranks what remains across every critical surface, so effort goes where an attack could actually reach. Neither depends on someone reading an alert in time.

There is a second, quieter reason. Everything downstream is priced against volume. Analyst hours, tooling tiers and managed services scoped to alert load all cost less when there is less to process. Prevention is the only investment that reduces the size of every other line item.

The analyst view has moved the same way. Gartner® forecasts that “By 2030, preemptive cybersecurity solutions will account for 50% of IT security spending, up from less than 5% in 2024, and replace traditional “stand-alone” detection and response solutions as the preferred approach to defend against cyberthreats.⁵

Action 1:

Make attack surface reduction and exposure management the foundation, and fund them first

2. See Every Critical Attack Surface, Not Just the Ones With CVEs

Most organizations have good visibility into one surface and partial visibility into the rest. Endpoint vulnerabilities are counted, scored, and reported. Alongside them sit misconfigurations, shadow IT and exposed external services, unnecessary access to legitimate tools, internal AI usage, and cloud workload misconfigurations. An attack chain crosses several of these, and a console built around one of them cannot show the chain.

Three of them carry most of the risk for a lean team. Abuse of the legitimate tools already installed appears in 84% of major attacks.⁶ Internal AI usage is growing faster than any process to govern it. Software vulnerabilities are, for the first time in the Verizon report's nineteen-year history, the most common way a breach begins.² Only a third produces a CVE.

The enterprise answer is a discovery product per surface, each with its own console and its own stream of findings. For a lean team, the test is simpler: one inventory, in the console the team already opens, covering the surfaces that carry risk rather than every surface a vendor can name.

Action 2:

Discover every critical attack surface, not only the ones that generate CVEs

3. Prioritize the Risks That Can Impact You

Vulnerability management is one of the few security controls almost every organization already runs, and the case for it is stronger now than ever. Over the period covered by the 2026 Verizon DBIR, only 26% of vulnerabilities in the Known Exploited Vulnerabilities (KEV) catalogue were fully remediated, down from 38%, and median time to patch rose from 32 days to 43.²

The practice was designed for a slower environment. It is bounded by the CVE list, while two of the three surfaces above produce no CVE at all. It is ordered by severity, which answers how bad a flaw is in the abstract rather than whether an attacker could reach that asset. And it is worked as a backlog, while volumes climb faster than any backlog clears.

Risk-based vulnerability management improves prioritization with information on which vulnerabilities are actively being exploited in the wild, AI capabilities that predict exploitability, asset criticality and threat intelligence. Exposure management is the broader step: misconfigurations, standing privilege, shadow IT, exposed external services, unnecessary access to legitimate tools and internal AI usage, all ranked by reachability and exploitability rather than score. Attack path analysis is the step that changes the work: it reveals which exposures chain into a route to something worth protecting, so a small team closes the few that complete a path instead of working down the many that do not.

Run as a repeatable cycle rather than a one-off project, Gartner's five-stage Continuous Threat Exposure Management (CTEM) is a good example that follows: scoping, discovery, prioritization, validation and mobilization. Often the stage that fails is the last one. Finding exposures is comparatively easy but addressing them before automated attacks can exploit them remains a challenge for many organizations.

Action 3:

Rank exposure by what an attacker can actually reach,
not by severity score alone

4. Harden Around Real Behaviour, and Let It Adapt

Attack surface reduction means removing the tools, privileges and paths a given user does not need, so that abuse of legitimate software has less to work with. Given that 84% of major attacks lean on tools already present, this is where prevention pays back fastest.

The test that separates a control that survives from one that quietly stops working: if it needs a periodic manual review to stay accurate, and there is no reliable hour for that review, it will be out of date within two quarters. Ask vendors what their control does on its own, three months in, with nobody watching.

The reason hardening has to be tailored is that a blanket rule set fails in both directions. PowerShell sits on every endpoint and a handful of administrators actually use it. Remove it for everyone and their work breaks; leave it for everyone and the tool attackers reach for first stays available across the estate.

The reason it has to be autonomous is maintenance. Environments and threats change continuously, and a rule set adjusted by hand at a quarterly review is out of date before the next one arrives. Hardening that adapts on its own, to how each person actually works and to how threats move, removes that overhead. It is the barrier professionals name first: 38.0% point to the effort of maintaining hardening rules and exceptions.¹

Action 4:

Harden around how each user actually works, and choose an approach that adapts rather than one that needs maintaining

5. Consolidate Rather Than Add

Exposure data belongs alongside the endpoint, identity and network signals your team already reviews, in the same console. A separate product with its own agent and its own learning curve is the version that quietly stops being used.

Two practical constraints follow. Start with the two links from earlier, unnecessary access to legitimate tools and ungoverned internal AI usage, because in a mid-market environment both are finite lists. And augment before you replace: exposure management and adaptive hardening can run alongside an existing endpoint or EDR tool, which removes the largest single objection to starting at all.

Action 5:

Consolidate into the console you already run,
and start with the two surfaces that have no CVE

Prevention First, Not Prevention Only

For a lean team this breadth can feel out of reach, and the feeling has a cause. The constraints chain: there are no spare hours to operationalize what you already own, so gaps persist, so another tool gets bought, and you end up again, with fewer hours. The sequence below is built to break that chain rather than add to it.

None of this argues for less detection and response. Something always gets through, and when it does you want it seen and stopped quickly, at hours when nobody is at a desk. Nearly half of teams say they cannot cover a full rotation,¹ and 95% of practitioners report a skills gap on their own team,⁷ so for most mid-market organizations that is a service rather than a hiring plan.

The argument is about order. Reduce the surface first, because that layer does not degrade as volume rises and it lowers the cost of everything after it. Then accelerate protection, detection, and response with AI and consider AI-enabled MDR services to help you gain speed in responding to attacks without adding complexity.

Taken together, the six actions are a sequence a lean team can start this quarter without a new programme:

- ▶ Make attack surface reduction and exposure management the foundation, and fund them first.
- ▶ Discover every critical attack surface, not only the ones that generate CVEs.
- ▶ Rank exposure by what an attacker can actually reach, not by severity score alone.
- ▶ Harden around how each user actually works, and choose an approach that adapts.
- ▶ Consolidate into the console you already run, starting with the surfaces that have no CVE.

Where GravityZone Fits

Bitdefender GravityZone is built around this sequence. Its exposure management capabilities consolidate risk across endpoint vulnerabilities, misconfigurations, external attack surface and internal AI usage in one view, with attack path analysis ranking what an attacker could actually reach. GravityZone Proactive Hardening and Attack Surface Reduction (PHASR) delivers hardening that adapts to how each user genuinely works rather than a static rule set, and both are available as an add-on and as standalone capabilities alongside a third-party endpoint tool.

[↪ Explore Prevention-First Security with the Bitdefender Gravityzone platform](#)



Endnotes

- 1 Bitdefender 2026 Cybersecurity Assessment, a survey of 1,200 IT and cybersecurity professionals across Europe, North America and South-East Asia.
- 2 Verizon 2026 Data Breach Investigations Report, based on more than 31,000 security incidents and 22,000 confirmed breaches across 145 countries, covering November 2024 to October 2025. Remediation and time-to-patch figures reflect Tenable Research data contributed to the report; the AI analysis was conducted with Anthropic's safeguards team.
- 3 Center for Strategic and International Studies, "Beyond Autonomous Attacks: The Reality of AI-Enabled Cyber Threats," April 2026.
- 4 ENISA, "ENISA's view on Cybersecurity in the Frontier AI Era," European Union Agency for Cybersecurity, 7 July 2026.
- 5 Gartner, Build Preemptive Security to Avert Weaponized AI Risks, Emerging Tech and Trends Security Research Team, 10 September 2026. GARTNER is a trademark of Gartner, Inc. and/or its affiliates.
- 6 Bitdefender Labs research into 700,000 security incidents.
- 7 ISC2 2025 Cybersecurity Workforce Study, 16,029 respondents, December 2025.
- 8 IBM Institute for Business Value, "Capturing the cybersecurity dividend," January 2025. Conducted with Oxford Economics and published in partnership with Palo Alto Networks; 1,000 executives across 21 industries and 18 countries.
- 9 Cyber Security Agency of Singapore, "Advisory on Risks associated with Frontier AI Models," AD-2026-004, 15 April 2026.
- 10 Bitdefender, "Technical Advisory: OpenClaw Exploitation in Enterprise Networks," Bitdefender Business Insights, 5 February 2026.

Bitdefender is a cybersecurity leader delivering best-in-class threat prevention, detection, and response solutions worldwide. Guardian over millions of consumer, enterprise, and government environments, Bitdefender is one of the industry's most trusted experts for eliminating threats, protecting privacy, digital identity and data, and enabling cyber resilience. With deep investments in research and development, Bitdefender Labs discovers hundreds of new threats each minute and validates billions of threat queries daily. The company has pioneered breakthrough innovations in antimalware, IoT security, behavioral analytics, and artificial intelligence and its technology is licensed by more than 200 of the world's most recognized technology brands. Founded in 2001, Bitdefender has customers in 170+ countries with offices around the world.

Release Date: September 2026

For more information, visit <https://www.bitdefender.com>.

Romania HQ
Orhideea Towers
15A Orhideelor Road,
6th District,
Bucharest 060071
T: +40 21 4412452

US HQ
111 W. Houston Street,
Suite 2105, Frost
Tower Building,
San Antonio, Texas
78205, USA