SANS | Research Program

WHITE PAPER

The Exposure Gap: From Vulnerability Management to AI-Driven Attack Surface Control

Written by **Chris Dale**

June 2026

seemplicity

©2026 SANS™ Institute

Introduction

What used to be castle walls and thoroughly inspecting traffic in and out has expanded to multicloud, hybrid environments, operational technology mixed with IT, off-prem identity systems, and third-party ecosystems that change daily. Traditional vulnerability management was not designed for this diversity and scale.

When this is combined with the two facts of cybersecurity—vulnerabilities increase steadily over time, and attackers are weaponizing and exploiting vulnerabilities faster than ever—it no longer takes expert talent to be an attacker, just an AI model and ideas of mischief.

To put the problem into perspective: More than 48,000 common vulnerabilities and exposures (CVEs) were published in 2025 (see Figure 1). Most organizations realistically can't remediate them all. The Exploit Prediction Scoring System (EPSS) provides organizations more insights into vulnerabilities that could be exploited, but it is not enough and it covers only parts of the overall complexity.

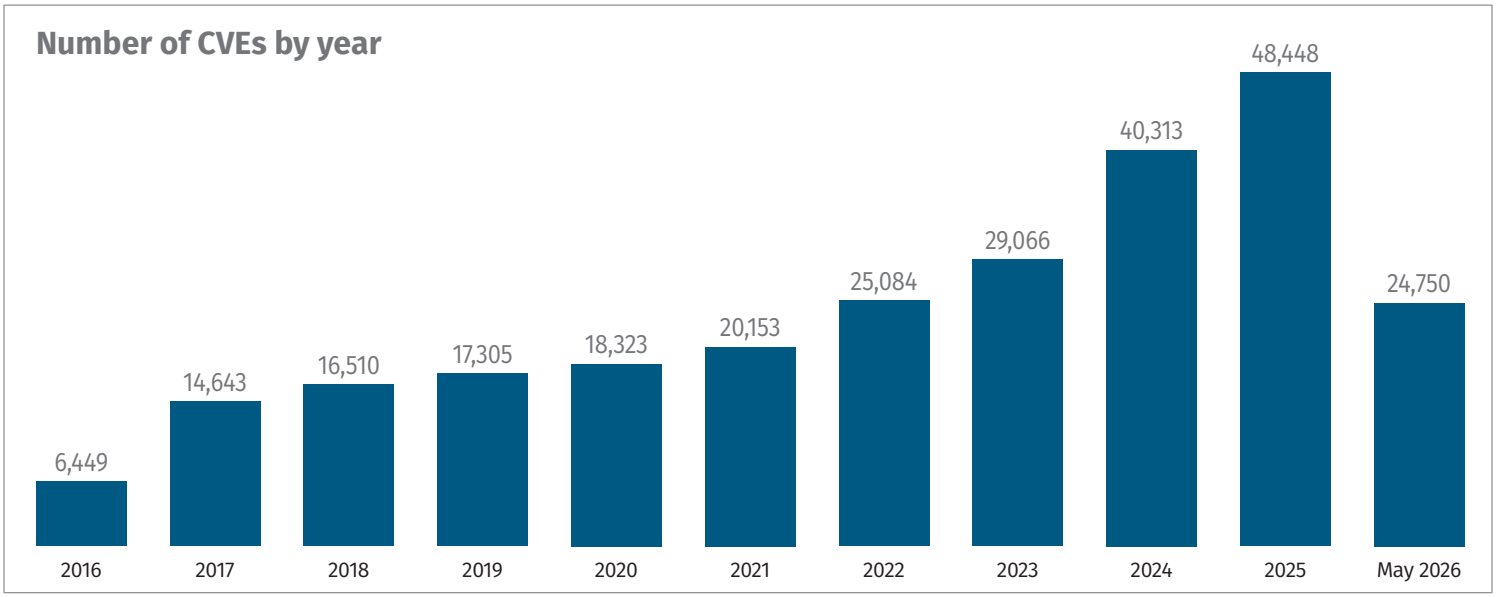

Figure 1. CVE Publication Growth from 2016 to May 2026[1]

In addition, vulnerabilities are not the only vector: Supply chains are being poisoned, attacking organizations from the inside out with new and novel attacks against package managers and code dependencies.

Cybersecurity evolves quickly. Developers and the "go-to-market" situation force organizations to *move quickly*. At the same time, the number of vulnerabilities and how attackers are using them is changing quickly. That's why it's important for defenders to also move quickly, and pitching AI against AI is, at this point, one of the better ways to start fighting back.

CVSS and EPSS Definitions

The **Common Vulnerability Scoring System (CVSS)** is a standardized framework for scoring the severity of vulnerabilities based on defined metrics.

The **Exploit Prediction Scoring System (EPSS)** is a model that tries to predict the likelihood that a vulnerability will be exploited in the wild.

[1] CVEdetails, "Vulnerabilities by Date," www.cvedetails.com/browse-by-date.php

The Limits of Traditional Vulnerability and Exposure Management

We want to explore AI solutions, but before we do, we need to explore why current approaches are structurally insufficient, not just under-resourced.

Vulnerability management is the capability to identify and remediate risk within the organization, ideally *before* that risk can impact the business. One of the key challenges, however, is that vulnerability management is typically distributed across a wide range of services and environments.

Consider your own environment. You likely have:

- **On-premises networks and applications—**These require security scanning across infrastructure and workloads.
- **Cloud environments—**Like most organizations, you have likely adopted some services in the cloud. Cloud security tools help identify risks and misconfigurations in those environments.
- **Endpoints—**Users carry out their work on devices such as Windows or Mac computers, which host a wide range of applications, including web browsers and the notoriously risky browser extensions.
- **Code repositories—**If you develop software, you almost certainly have code repositories. These should be scanned for flaws and vulnerabilities so threats can be identified earlier in the development life cycle.

There are also many other areas to consider, including identity, SaaS, API security, containers, and Kubernetes. It all depends on the organization. There is no one-size-fits-all approach. In our experience, most organizations have a *modern hybrid attack surface*, relying on a broad mix of security solutions that varies significantly from one organization to another.

Visibility between these utilities is often siloed and worked on independently, not as a whole from an organization's risk point of view. Furthermore, each solution produces individual lists of assets and software bills of materials, and has its own understanding of risk.

With risk prioritization comes a plethora of different approaches. Perhaps your team is bogged down in fixing a 9.8 CVSS vulnerability (a critical one), but it's on an isolated system not connected to the internet, all while a weaponized vulnerability on an internet-exposed asset is waiting in the queue. That brings us to the next point: Point-in-time scans miss ephemeral assets and drift in configuration between scan cycles.

It is important to recognize that most traditional vulnerability management is a point-in-time snapshot of the actual risk present. This means that the system's data is stale immediately. The classic question is: *How often do you scan?* A better approach would be to look at the system from a continuous point of view.

Where AI Changes the Calculus

The AI revolution has propelled attackers into previously unimaginable dimensions of speed, scale, and creativity. AI has tremendously increased the speed at which attackers can go from discovering vulnerabilities to exploiting them. Time-to-exploit (TTE) measures the gap between CVE disclosure and confirmed exploitation, and it is down to just 20 hours! (See Figure 2.) This is fast … but is it too fast?

Models like Anthropic's Mythos suggest a future where both vulnerability discovery and patch development happen at an even greater pace. Specialized large language models (LLMs), allegedly capable of outperforming professional security researchers and uncovering large numbers of vulnerabilities in code that has been scrutinized for years, are likely to shake up the security landscape even further.

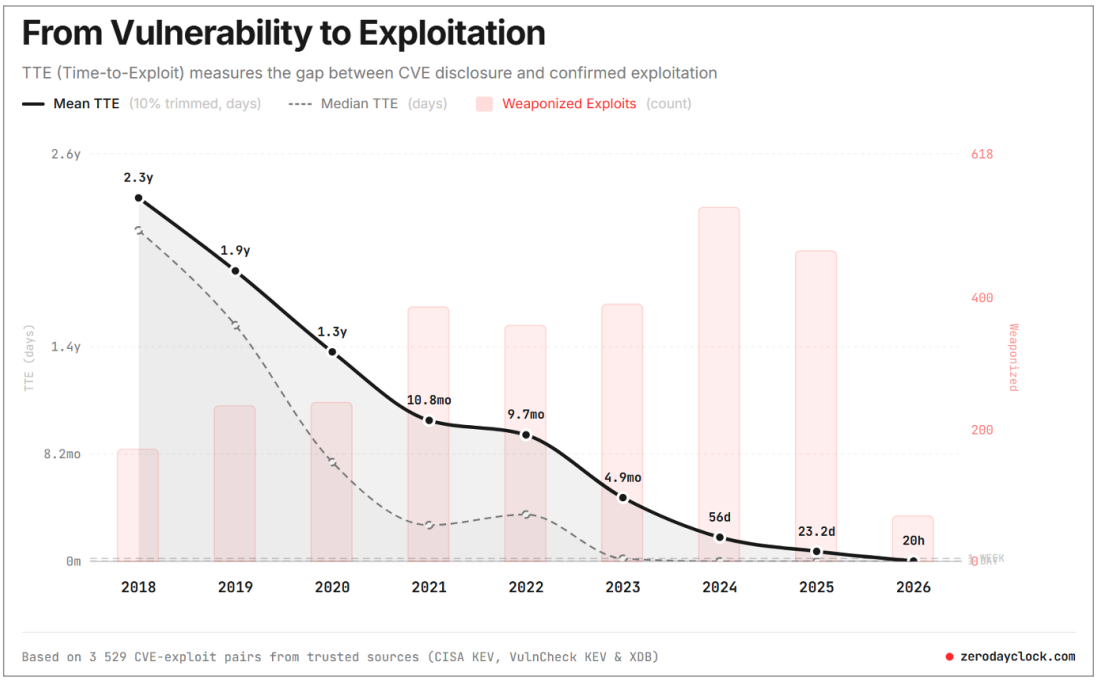

Figure 2. Timeline Showing That Vulnerabilities Are Increasingly Faster to Weaponize[2]

One of the first steps defenders can take toward a more manageable future is data consolidation and normalization. Storing data across multiple systems leads to a loss of valuable context. Key contextual datapoints for analysts include factors such as business value, whether an asset is publicly exposed, whether it is accessible from internal networks, and other factors. This data could be stored within an asset inventory system, now often defined as an attack surface management solution.

Attackers are not the only ones who can leverage the advancements of AI. When we tack on AI, we can get a lot of quick wins. LLMs, for example, are a great way to build out context about environments where infrastructure is ephemeral by design. Containers spin up and die in minutes, serverless functions execute and disappear, and temporary test environments quietly introduce new exposure paths. Traditional asset management simply can't keep pace here, but machine language–driven classification can identify patterns, recognizing when a short-lived resource behaves like a production system, exposes sensitive data, or introduces risk.

[2] Zero Day Clock, https://zerodayclock.com

Another valuable use for AI is to help contextualize even further. It can help prioritize which vulnerabilities and challenges, from your different sources, need to be remediated first. It also can offer helpful remediation advice that will sever the routes to your crown jewels or prioritize quick fixes across the environment to minimize risk.

Exploitability, in the context of vulnerability intelligence, is becoming increasingly important given the sheer number of vulnerabilities surfacing at any given time. Receiving up-to-date information about when an older vulnerability becomes a known exploitable vulnerability (KEV) is a clear example of effective vulnerability intelligence. Likewise, insight into whether a public proof-of-concept exists, and whether a patch is available for deployment, are equally critical components.

Ultimately, defenders are being forced to operate at the same speed and scale as their adversaries. The combination of consolidated data, enriched context, AI-driven prioritization, and real-time exploitability intelligence is no longer a luxury, but a necessity. Security teams that fail to adapt will drown in noise, while those who successfully operationalize these capabilities can shift from reactive firefighting to proactive risk reduction. In a landscape where exploitation can happen within hours, the real advantage is not just visibility; it's the ability to understand, prioritize, and act faster than the attacker.

Without Visibility, the Fight Is Already Over

Imagine that organizations today are in a constant struggle between shipping products quickly and ensuring security can keep up. Now add AI to the mix, and suddenly every person in your organization—not just the engineering team—has access to an intern who can code, develop products, and, with a single prompt, set up tunnels from the internet straight into the heart of your business. All of this can happen without a real understanding of what is being built, just a desire to solve problems quickly, bypassing firewalls in the process.

We call this vibe coding, and it produces what we describe as "vibe ware." What used to be a significant challenge with shadow IT has now grown far worse with the advent of AI. It is not necessarily all bad. The constant innovation is likely beneficial for the organization in the long run. But without the necessary governance, the risk could be too great.

The author of the paper has seen this happen firsthand. A vibe-coded application needed to receive incoming requests from an integrated third party, and Claude Code's solution to the problem was to, with a few keypresses, expose the application to the internet via Cloudflare. No account signup, no configuration, just a single terminal with a vibe coder behind the keyboard. The employee had exposed the internal network to the outside, and within 10 seconds of exposure, attackers were already trying to exploit it.

Visibility can be gained, but visibility isn't a single tool. It's the cumulative picture of assets, identities, traffic, and behavior across your entire environment. We need to ensure we can see all devices and applications on the network, attached to the cloud, in use in workstations, and across the enterprise. When visibility is gained (e.g., by consolidating the data from different tools and platforms into one unified platform), we need to stop thinking in terms of patch counts and scan coverage. Instead of thinking in lists, we need to start thinking in terms of graphs—after all that's what attackers do. They ask themselves, *How do I get to the crown jewels?* AI really shines here. Its extensive knowledge can apply attacker-specific tradecraft, in a defensive setting, enabling the most impactful vulnerabilities to be fixed first. Attackers think in graphs, while defenders think in lists. With the help of AI, we can have a side-by-side comparison of a traditional CVSS-ranked patch list vs. an AI-prioritized list ranked by blast radius.

Without continuous discovery and governance of assets, efforts will be futile. It is imperative to maintain a constantly evolving inventory of what needs to be defended, especially in the AI era we are now living in. In fact, one of the most effective ways to defend ourselves in this revolution is to have nothing to defend.

How to Have Nothing to Defend

Having nothing to defend sounds ideal, but how do we get there? By contextualizing what is exposed, identifying where gaps exist, and understanding the paths to their crown jewels, organizations can move beyond simply patching software and begin shifting architecture. This enables them to properly implement network segmentation, adopt zero-trust principles that validate users and devices before granting access, and reduce unnecessary exposure across the environment.

One of the primary strategies for organizations, both now and in the future, is to significantly reduce exposure. This means removing paths to critical assets and limiting the blast radius of potential impact. When AI is given the right context, data, and objectives, it can drive meaningful improvements in architectural decisions rather than just identifying another vulnerability to patch.

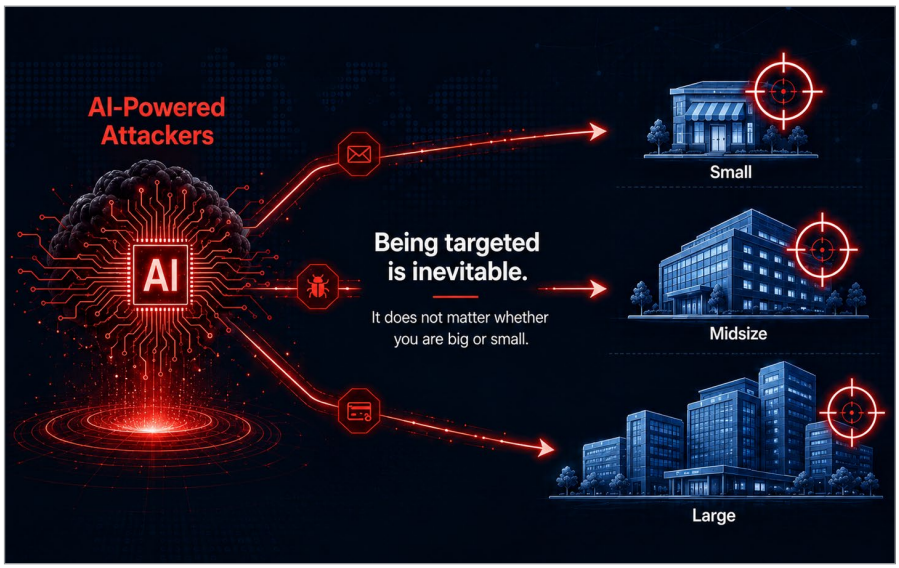

Figure 3. Attackers find vulnerabilities and can weaponize them against anyone.

Have no doubt, regardless of whether your organization uses AI defensively, it is already being used on the offensive, and you are a target (see Figure 3). It does not matter whether you are big or small, being targeted by attackers is inevitable. The real question is not, *Will we be hacked?* The answer to that is yes. The question you should be asking is, *Will we know in time to do something about it before damage is done?*

Defenders Are Driving Go-Karts, While Attackers Are in Formula 1

Attackers are at the point where fully automatic agentic frameworks can do everything from researching to exploiting and traversing your networks in real time. To counter this shift in balance, it is essential for defenders to leverage similar tools.

What makes this shift even more challenging is that it is not just about access to technology but about the governance gap between attackers and defenders. Attackers operate without constraints. For them, AI is the Wild West. There are no compliance requirements, no internal approvals, and no consequences if something breaks or behaves unpredictably. Defenders, on the other hand, must navigate regulations, risk tolerance, and organizational controls, especially in heavily regulated industries where the adoption of AI is approached with caution. This creates an uneven playing field where attackers can move fast and experiment freely, while defenders are often slowed down by necessary but limiting guardrails. The result is clear: This is no longer a balanced fight.

While attackers continue to operationalize and move faster, driving their mean TTE lower and lower, defenders must begin measuring themselves just as rigorously. It is not acceptable for attackers to be "driving Formula 1 while we are still in go-karts." Organizations need to take a clear stance on their mean time-to-resolution (TTR) metrics. How fast are we *actually* fixing issues?

More importantly, the focus should not be on fixing *everything*, but on fixing the *right* things. High-exploitability findings must take priority over sheer volume. If the exposure window is too large, and in some cases we are talking about hours, it becomes clear how difficult this challenge really is.

Metrics such as the average time a weaponized vulnerability exists in the environment before remediation are far more valuable than looking at patch cadence alone. A more dynamic approach to patching is required, one that is driven by contextual risk. This allows organizations to patch on demand, addressing what matters and where it matters, and doing so quickly.

When attackers traverse graph-based paths toward our crown jewels, we need to ask an important question: *Instead of giving them a racetrack for their Formula 1 car, why not force them into a parking lot maze?* In that kind of environment, even a go-kart would be faster (see Figure 4).

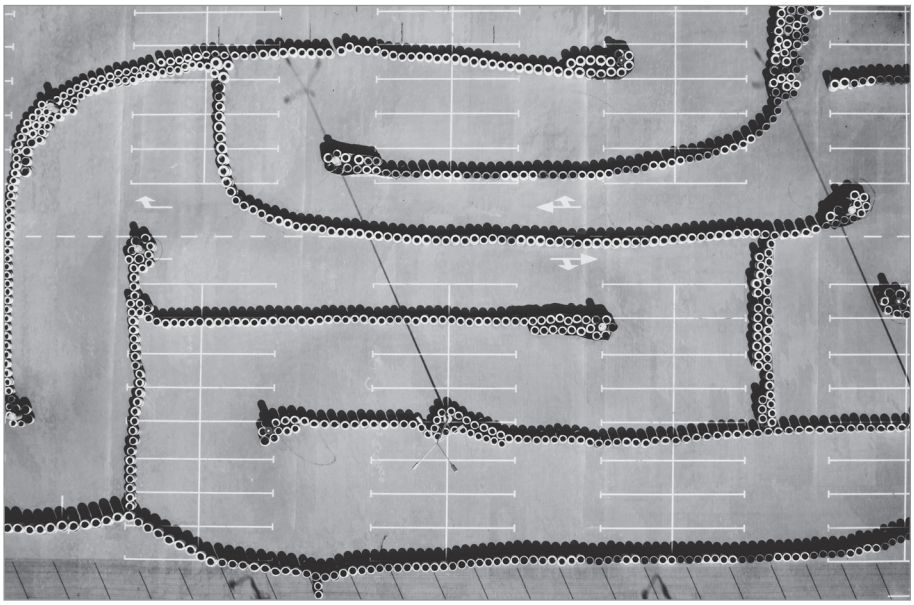

Figure 4. Changing the Parameters of the Race

But enough with the analogy. The real point is this: Organizations must focus on reducing critical attack paths to crown-jewel assets over time, not just on patch management.

Another important goal for organizations is not only to prevent attackers but also to detect threats as early as possible—ideally, long before they can reach any crown jewels. Attackers do not announce themselves. Instead, they rely on the gap between initial access and detection. Many breaches are discovered weeks or even months after the initial compromise, meaning the damage is often done long before any alert is triggered.

Hopefully, by gaining visibility, contextualization, and advice from AI that understands the fundamentals of hacking, organizations can work toward making environments hard to traverse and remediating risk before the impact is irrecoverable.

Taking this to the next level comes down to reporting and executive buy-in. Instead of technical teams reporting on how many issues were remediated last month or how many critical vulnerabilities were patched, we should be translating those activities into business risk reduction by answering questions such as, *How many attack paths were removed?* and *How are we challenging attackers on their own ground, rather than focusing on only one slice of exposure management, vulnerability management?* If AI can help us translate technical risk into business risk, while also providing context and recommendations for long-term strategic mitigation, then we are setting ourselves up for success.

Conclusions

AI does not solve exposure management. Instead, it removes the bottlenecks that have historically prevented human teams from effectively doing defensive programs like exposure management. The scale, speed, and complexity of modern environments have outgrown manual processes, fragmented tooling, and static prioritization models. AI changes the equation by enabling defenders to operate with the same tempo and context as attackers, but only if it is applied correctly.

It starts with data. AI will amplify whatever it is given. Poor data quality, fragmented inventories, and missing context will only lead to faster, but still flawed, decisions. Consolidation, normalization, and continuous discovery are therefore foundational. Without a reliable understanding of assets, exposure, and business context, even the most advanced models will fail to produce meaningful outcomes.

The next shift is in how we measure success. Traditional metrics, such as vulnerability counts, scan coverage, and patch volume, are no longer sufficient. They measure activity, not impact. The question is no longer how much to fix, but what to fix. *Where can we reduce the most risk with the least effort? Where do we break the attacker's path?* This is the core of modern exposure management.

Organizations must prioritize the reduction of attack paths over the remediation of individual vulnerabilities in isolation. Attackers do not think in lists, they think in graphs. Defenders must do the same. Reducing connectivity to crown-jewel assets, limiting blast radius, and eliminating high-risk pathways will have a far greater impact than chasing ever-growing vulnerability backlogs.

This also requires a shift in how risk is communicated. Executives do not need to know how many vulnerabilities were patched last month. They need to understand how risk has been reduced, how many attack paths were eliminated, and, as a result, how much harder it is for an attacker to reach critical assets. AI can play a key role here by translating technical findings into business-relevant insights, supported by clear and actionable recommendations.

Finally, the reality is that attackers are already leveraging agentic AI to automate reconnaissance, exploitation, and lateral movement. This is not a future concern; it is happening now. Defenders must respond in kind. AI is no longer optional. It is a necessary capability to keep pace, to prioritize effectively, and to act within shrinking exposure windows.

The organizations that succeed will not be those that fix the most vulnerabilities, but those that reduce the most risk.

Sponsor

SANS would like to thank this paper's sponsor:

seemplicity