

HOW IMPORTANT IS SPEED IN **VULNERABILITY MANAGEMENT?**

UNRAVELING CLAIMS
ABOUT TIME TO EXPLOIT



TABLE OF CONTENTS

INTRODUCTION 02

Purpose of this report

THE AVERAGE DEFENDER'S EXPERIENCE 04

So... CVEs?

How dangerous is CVE exploitation?

WHAT'S REALLY HAPPENING IN VULNERABILITY MANAGEMENT? 08

Question 1: What does CVE exploit activity look like over time?

Question 2: How is AI changing CVE exploitation?

PROTECTING AGAINST CVE EXPLOITATION (AND BIKE THEFT) 26

Key takeaways for defenders

Closing thoughts

UNDERSTAND & MITIGATE CVE EXPOSURE WITH YESWEHACK 32

For pre-emptive and emergency CVE mitigation



INTRODUCTION

“The average time to exploit (TTE) vulnerabilities is now negative.”

If you’re involved with vulnerability management, you’ve probably heard this. Precise figures vary, but one of the most widely cited is that TTE is now -7 days. This claim originates from Mandiant’s M-Trends 2026 report, which is based on analysis of incident investigations during 2025.

Presumably, security teams should think, *“Wow, if TTE is negative, we’ll have to patch **really** fast!”*

Except, that wouldn’t work, would it? If vulnerabilities were exploited en masse before anyone knew about them, vulnerability management (at least, as it currently exists) would be a waste of time. It would be the definition of closing the barn door after the horse has bolted. Even the mass replacement of patching with faster mitigations would be too little, too late.

Of course, a negative TTE doesn’t mean a vulnerability will be exploited *everywhere* before it’s published as a CVE. Just that exploit activity begins pre-publication. If that point were made a little more clearly, a little more often, this report would probably never have happened.

Instead, most of the time, when you hear, *“TTE is now negative”*, it’s immediately followed by *“...and attackers use AI now, so they can exploit CVEs before patches are available.”*

If we accept these two statements uncritically, they might lead us to conclude:

1. Vulnerabilities are exploited hours or days before you know about them
2. Exploitation happens en masse before you can respond
3. This is the *typical experience* for organisations and their security teams

But is this true? Let’s find out.

PURPOSE OF THIS REPORT

We're not disputing TTE findings by Mandiant or anyone else. Instead, we are questioning whether claims made across the industry about TTE and AI use in CVE discovery and exploitation tell the full story about the current state of vulnerability management.

The purpose of this report is to:

- Examine the available data on CVEs and exploit activity
- Investigate how AI may be affecting CVE discovery and exploitation
- Attempt to understand the *typical experience* of defenders
- Identify potential learning points for defenders

Note: This report focuses on exploitation of CVEs specifically, **not** the wider field of exposure management. Also, we will endeavour to distinguish between what the data (from various sources) actually says, and our interpretation of it.



✂ THE AVERAGE DEFENDER'S EXPERIENCE

Imagine you're a cyclist in the 1950s. You cycle everywhere, and dutifully lock up your bike whenever it's unattended using the tried-and-true method of chain-plus-padlock.

Then, some pesky inventor comes out with a cordless angle grinder, which could separate you from your beloved bike in a fraction of the time it would take to cut the chain by hand.

What do you do? In all likelihood, you do nothing.

Why? First, because unless you work in construction, you likely don't hear about the invention. Further, cordless grinders wouldn't be widespread for at least another decade, and it would be considerably longer before they were readily available and affordable for individuals.

You *might* have your pride and joy stolen by a grinder-wielding bike larcenist. But if you did, you might reasonably feel aggrieved at your bad luck. In practice, it was an unreasonable threat to anticipate, and you'd be very unfortunate to fall victim to it.

Bike theft was nothing new. Chain cutting could be foreseen. But the specific risk posed by cordless grinders was not something you could reasonably have predicted.

If it were possible to map bike thefts against the mass availability of cordless grinders over time, we'd probably see a small number of cases initially, a gradual increase, and then a substantial rise as the tactic became widely known. We'd also see an influx of grinder-resistant security products entering the market in response.

There's more to this, of course, and we'll return to it later.

SO... CVES?

What does any of this have to do with CVEs? Quite a bit, actually.

Negative TTE means attackers exploit vulnerabilities before they are published as CVEs. This is often because the CVEs in question are **zero days**, so-called because software vendors have no time to develop a patch prior to the vulnerability becoming relevant.¹ If TTE has fallen into negative territory, we might expect to see a higher proportion of zero days among new, exploited CVEs compared to previous years... and this appears to be true.²

Which sounds scary, until you really think about it. Does it seem likely that attackers are exploiting CVEs **in significant numbers** before they are disclosed? Wouldn't the exploit activity itself prompt investigation that would result in a CVE being published?

Well, yes. That's exactly how it works. In fact, there are cases where a vulnerability is exploited months or years before it becomes a CVE. This is possible *only because* the initial exploit activity is very limited in scope, and not enough to prompt a thorough investigation.

Sources in the industry have variously stated that TTE is now anywhere from -7 days³ to a few hours. There's plenty of data to back up these claims, but one fact is rarely mentioned: these figures represent **Time To FIRST Exploit (TTFE⁴)**, *not* time to mass exploitation. In our view, a raw TTE/TTFE figure is not particularly useful for understanding a typical organisation's real-world exposure.

The IT security industry is fortunate to have some outstanding research resources. Unfortunately, the industry at large also has a tendency to take evidence-based data points, add some dubious claims, and conclude the sky is falling.

1. "Relevance" in this case could be due to exploit activity, availability of an exploit kit, public awareness, etc.
2. Actually, there is debate over this, depending on how we define zero day vulnerabilities
3. [M-Trends 2026](#), Mandiant
4. Or even Time To First Known Exploit (TTFKE)

A more helpful approach would be to look at how exploit activity is spread over time, whether that's changing, and how quickly organisations would need to respond (once able) to minimise risk. Ideally, we'd draw some sensible conclusions from real-world exploit activity that will help defenders deal with the rising threat of CVE exploitation.

HOW DANGEROUS IS CVE EXPLOITATION?

The IT security community has accepted (at least through its actions) that CVEs are very much worth worrying about. Still, it's worth understanding the problem before we invest too much energy in potential solutions. For that, we'll lean on Verizon's Data Breach Investigations Report (DBIR).

The DBIR has never claimed to be a perfect representation of security incidents and breach activity. It's also difficult to compare the data across years, because the collection process has evolved (and grown) over time. Still, incomplete data is infinitely preferable to no data.

The 2026 DBIR found that "Exploit Vulnerability" (which relates specifically to CVEs, rather than the wider field of exposure management) was the most common initial access vector in data breaches. In fact, 31% of data breaches in the 2026 DBIR dataset⁵ are attributed to it, almost twice as many as phishing (16%), the next most common vector.⁶

CVE exploitation is also among the most dangerous actions in cyberattacks: around 98% of incidents in the 2026 DBIR dataset had confirmed data exposure.⁷

5. M-Trends 2026 also found CVE exploits were the most common initial infection vector in IR investigations at 32%

6. These figures are slightly skewed by a small number of widely-exploited CVEs. Still, CVE exploitation was present in 18% of data breaches in the previous year, and it remains to be seen if it will continue to reign supreme in the coming years

7. This type of calculation is highly susceptible to bias in the dataset, so take this with a pinch of salt

It's also worth noting that "System Intrusion", the broader DBIR category that CVE exploitation sits within, is by far the most common pattern in both security incidents and breaches. It has been the top pattern in breaches since 2022, and in incidents since 2024. So, while CVE exploitation is undoubtedly a massive issue for defenders, it should be noted that non-CVE-based technical attacks (e.g., malware, backdoors, etc.) represent at least the same degree of danger, and potentially more.

Nevertheless, CVE exploitation is undoubtedly a real and present threat for organisations, and one that seems to be growing over time.



✂ WHAT'S REALLY HAPPENING IN VULNERABILITY MANAGEMENT?



For this investigation, we relied on data from various sources. Unfortunately, that data is collected from different sources and using different methodologies, and in many cases the raw data is unavailable. That makes it practically impossible to normalise, and makes it difficult to draw definitive conclusions.

With that said, we do believe there are some valuable learning points along the way.

We're aiming to answer two questions:

- What does CVE exploit activity look like over time?
- How is AI changing CVE exploitation?

We've selected data sources that (hopefully) give us the best chance to understand a typical organisation's experience, so we can draw some useful conclusions.

QUESTION 1: WHAT DOES CVE EXPLOIT ACTIVITY LOOK LIKE OVER TIME?

Here we have a wealth of data to work with, coming variously from the DBIR, Zero Day Clock, DScope, and a variety of academic papers (several using internet telescopes similar to DScope).

Note that we're not arguing with the claim that, on average, TTE/TTFE is negative. We're questioning the subtext: the idea that CVE exploitation is an "everything everywhere all at once" situation, rather than (as we might reasonably expect) distributed over time.

❗ NEW ISN'T EVERYTHING

The first thing to consider is whether new CVEs are the greatest source of risk.

The 2026 DBIR notes that, as of February 2026, there were 1,526 CVEs listed in the CISA KEV Catalog, of which 991 had some exploit activity during the previous year.

If we look at the KEV Catalog, we can see that during that year-long period, 249 CVEs were added. Of those, 172 were new CVEs, while 77 had been published in previous years.⁸ So, new CVEs accounted for 69% of CISA KEV additions, but just 17.4% of all exploited CVEs. This begins to paint a picture.

Across all exploited CVEs in the period, 48% were deemed to have “persistent” exploit activity, meaning associated exploit activity was detected an average of 96% of days during that year. And of those persistently exploited CVEs, 80% were published in the CVE database **before 2024**, meaning defenders had, in theory, quite a bit of time to remediate them.

So not only did new CVEs represent a relatively small proportion of exploited CVEs, they were also a tiny proportion of *repeatedly* exploited CVEs during the period. Now, this is not a completely fair comparison, because only new CVEs from the very beginning of the period could conceivably have been observed so consistently throughout it... but this begins to suggest that TTE for new CVEs may not be the full picture.

Expanding on this point, Greynoise released a report in 2025⁹ specifically on what it calls “resurgent” vulnerabilities. The report notes that CVEs are sometimes exploited years after disclosure, and that exploit activity over time varies significantly, posing an extra challenge for defenders.¹⁰

✂ TRACKING CVE EXPLOIT ACTIVITY

Cybercriminals continue to exploit old CVEs because it *works*. It doesn’t matter when a CVE was published, only whether it’s present and exploitable in your environment.

Still, knowing exploit activity is spread over time is one thing. Understanding what that looks like in practice is another.

8. Note that you can’t assume when a CVE was published based on its identifier. For example, CVE-2023-52163 (Digiever DS-2105 Pro Missing Authorization Vulnerability) was published in March 2025 and added to CISA KEV in December 2025

9. “A Blindspot in Cyber Defence”, Greynoise, 2025

10. Greynoise also categorises some vulnerabilities as eternal, not resurgent, because they never went away



For this, we're first going to lean on DScope¹¹, an internet telescope that continually starts, runs, and terminates Amazon EC2 instances in every region and availability zone globally. It can be used to measure adversarial behaviour (including exploit activity) targeting cloud infrastructure.

DScope was also used by researchers at the University of Wisconsin-Madison to produce a 2023 report titled 'The CVE Wayback Machine',¹² which looked at real-world exploitation timelines for 63 CVEs. Combined with data from other sources, this allowed for the production of full birth-to-death activity models for these vulnerabilities.

The report had some interesting findings:

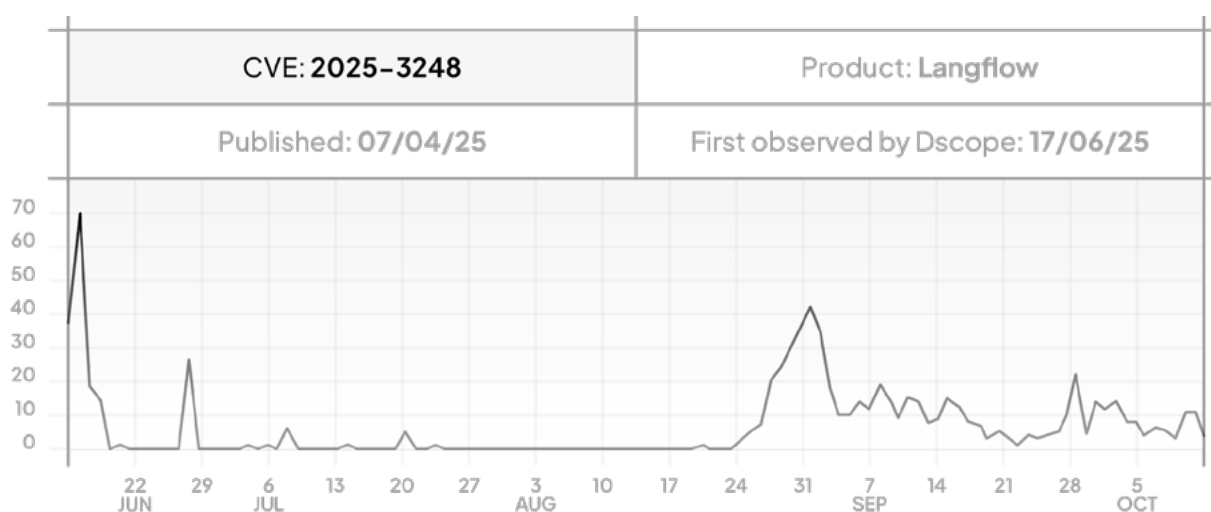
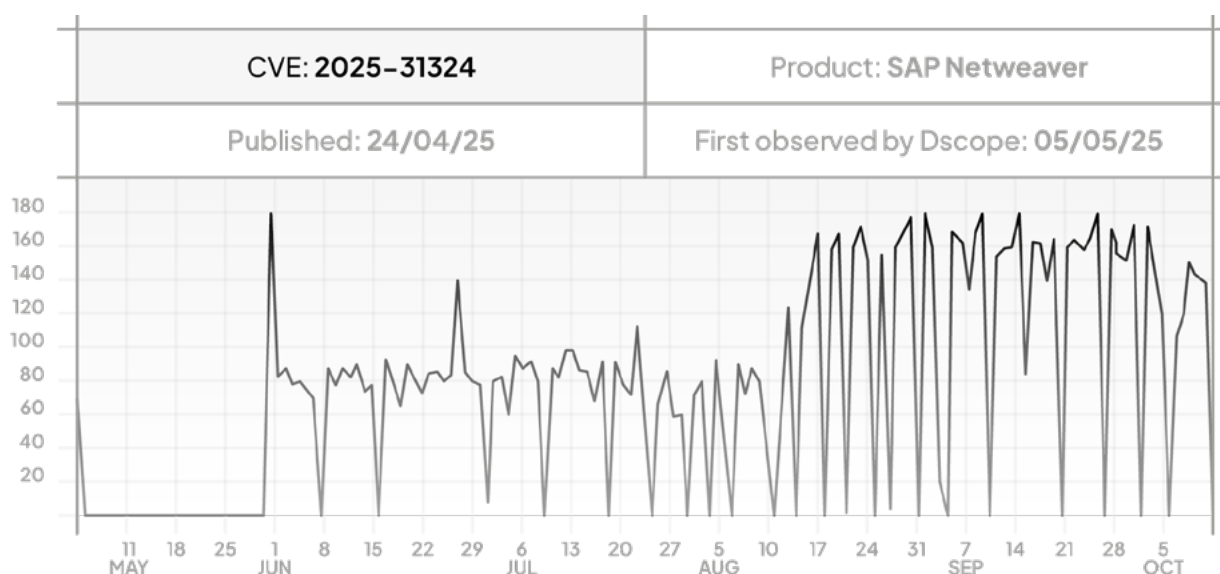
- "Exploit of critical vulnerabilities can follow very closely after announcement."
- "When attacks precede defenses [...] they often do so by a very brief period (only a few days)."
- "CVEs are primarily targeted in the days and weeks after publication, though others see sustained targeting over time"

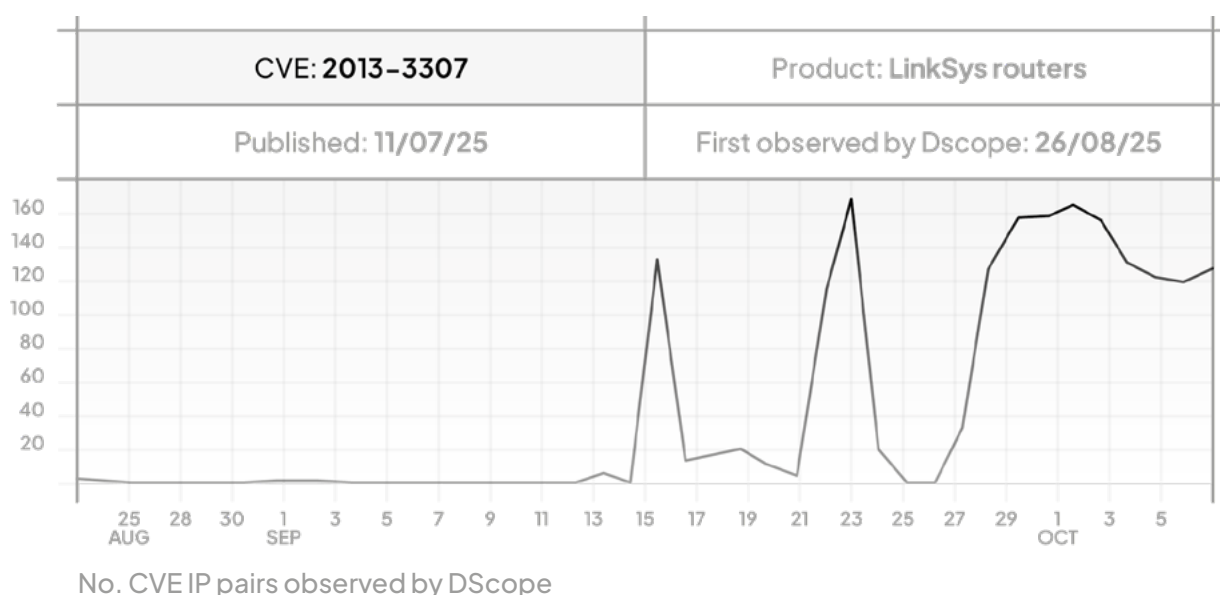
¹¹. You can learn more by reading "[DScope: A Cloud-Native Internet Telescope](#)" (E Pauley et al, 2023), or visiting dscope.org

¹². "[The CVE Wayback Machine](#)", E Pauley et al, 2023

The report also notes that in cases where attacks preceded CVE publication, it was often significantly before, but in small volumes. Again, the vast majority of exploit activity happens after publication.

Of course, 2023 is a long time ago in security terms. So we'll now focus on what the underlying tool, DScope, shows in recent history. Here it is observing three CVEs that were newly published in 2025:





Notice that in each case, the major spikes in exploit activity detected by DScope happen some weeks after publication. This doesn't mean exploit activity was nonexistent before then, but it does suggest activity was on a small enough scale to evade detection by an internet telescope.

Now, three CVEs are far from a representative sample. However, if you look at the DScope activity data more broadly, you'll notice that this trend is mirrored throughout.

Then there was the global **React2Shell** phenomenon. Disclosed on 3rd December 2025 (along with a vendor advisory and patch) exploit activity for CVE-2025-55182 was observed within hours.

Still, research shared by CloudFlare¹³ found that exploit activity began to trend significantly upward around 9th December (five days later) and continued through the following days, weeks, and months. In fact, React2Shell exploit activity is still observed today, so it remains a current threat.

Mirroring these findings, researchers at the CSIR Fourth Paradigm Institute in Bengaluru published a paper¹⁴ in March 2026 that broke down React2Shell exploit activity during December 2025.

13. ["React2Shell and related RSC vulnerabilities threat brief"](#), CloudFlare, December 2025

14. ["Internet-Scale Measurement of React2Shell Exploitation Using an Active Network Telescope"](#), A Singh et al, March 2026

Again, while exploit activity was observed within hours of CVE publication, the first spike in activity occurred five days after publication, and a high proportion occurred between weeks two and four.

KEY TAKEAWAYS

Add all this together, and what do we get?

1. Significant exploit activity typically happens within days or weeks of CVE disclosure.
2. Exploit activity frequently occurs over a considerable period (months or years).
3. Some exploit activity can occur prior to (or concurrent with) disclosure, but not much.
4. CVE resurgence is common, so exploit risk doesn't disappear after the initial flurry.
5. Old CVEs that weren't previously exploited may suddenly become active for the first time.

In short, while average TT(F)E is negative, this is *not* the typical experience for defenders. In fact, one of the major takeaways from the "CVE Wayback Machine" paper mentioned earlier is that Coordinated Vulnerability Disclosure (and defenders' actions based on it) is more effective than it has generally been credited with in recent years.

That said, it would be foolish not to concede that vulnerability management is getting harder for defenders, and that prompt remediation of exploitable vulnerabilities is more important than ever.

If you had fully remediated a typical 2025 CVE within a few days, you would have been protected against the majority of exploit activity. Of course, that wouldn't be much consolation if you'd already been breached during the initial flurry.

This is the dichotomy between risk calculations and the real world. Risk aims to score and prioritise threats for remediation and protective purposes. But real-world incidents are binary; they either happen, or they don't. Still, the faster you respond, the lower the chance you'll be an unlucky outlier.

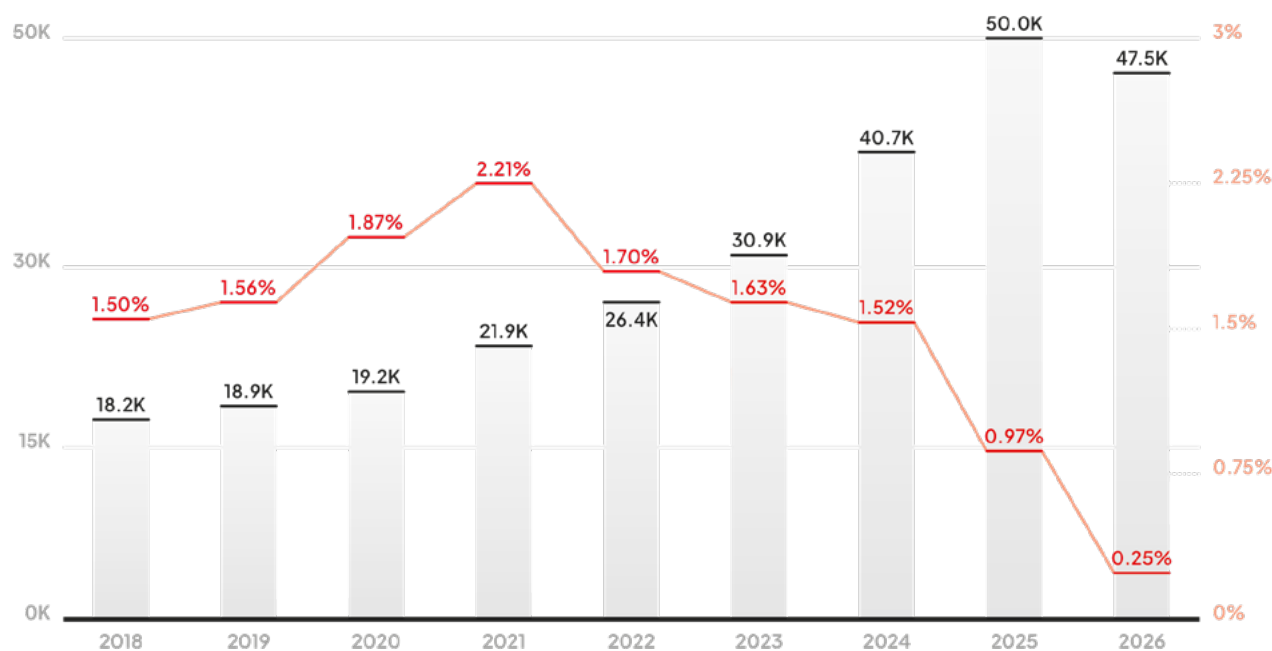
❗ WHY TTE ISN'T THE FULL STORY

One very interesting fact that gets a lot less attention than TTE is the number of newly published CVEs that are exploited. After peaking at 620 in 2024 (1.52% of CVEs published that year), this number has been in freefall ever since. In 2025, 483 new CVEs were exploited (0.97%), and so far in 2026 the number is just 118 (0.25%).¹⁵

Exploit Rate

Percentage of all published CVEs that are confirmed exploited in the wild, versus total CVE volume

— Exploit Rate (% of all CVEs) — Total CVEs Published (count)



Source: zerodayclock.com

Assuming a similar rate of CVE publication and exploitation through the rest of 2026, we might expect to see somewhere north of 71,000 CVEs, but only ~177 exploited.¹⁶

At the same time, ZDC identifies a huge rise in the rate of zero day vulnerabilities among newly published CVEs. Rising from 49.7% to 53.6% between 2004 and 2025, the rate has soared to 84.8% in 2026 so far.

¹⁵. Figures from Zero Day Clock (zerodayclock.com). There are minor discrepancies compared to CISA KEV, because ZDC incorporates additional data sources

¹⁶. Of course, this number may rise as 2026 CVEs are exploited for the first time in future years

ZDC lists TTE for 2026 CVEs as -23 hours, meaning exploit activity is observed and confirmed almost a full day prior to CVE publication. In these cases, the security community (and the affected vendors) can only respond once exploit activity is observed.

Two things to note here:

1. As we've seen, even when exploitation precedes CVE disclosure, *most* exploit activity will occur shortly afterward
2. A zero day rate of 84.8% sounds scary, until you realise that *most* of the CVEs exploited in 2026 were published in previous years

Unfortunately, the relatively small number of exploited 2026 CVEs doesn't equate to a significantly reduced workload for defenders. As previously noted, CVEs are often targeted persistently, so those from previous years may still be a concern, particularly if your infrastructure changes or grows.¹⁷

As of mid-August, some 182 CVEs have been added to CISA KEV in 2026, but 72 (39.6%) of those were published in previous years. In fact, 38 (20.9%) were published in 2024 or earlier. For these older CVEs, TTE will naturally be at least positive, and frequently measurable in months or years.

CVEs added to CISA KEV in 2026 so far

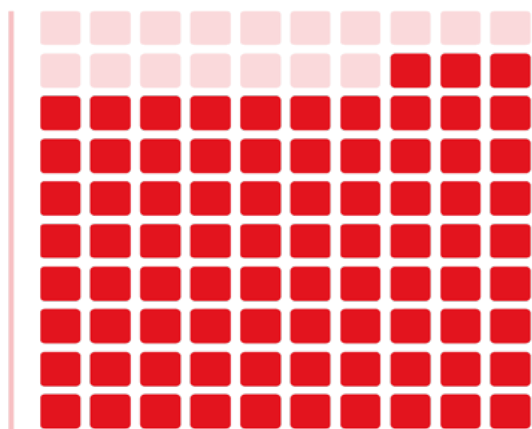


And this doesn't include any resurgent vulnerabilities, such as CVE-2023-47565¹⁸, added to CISA KEV in years past, but which remain relevant. If we look at 2025, where we have a full year's data to consider, "old" CVEs accounted for 30.9% of new CISA KEV entries, and a whopping 82.6% of CVEs with observed exploit activity.

¹⁷. For example, M&A activity is a common way to "inherit" CVE risk

¹⁸. QNAP VioStor NVR OS Command Injection Vulnerability

All CVEs exploited in 2025



82.6%

published in 2024 or earlier

CVEs added to CISA KEV in 2025



30.9%

published in 2024 or earlier

Notice that the proportion of new CVEs added to CISA KEV has actually fallen in 2026 compared to last year. Given the comparatively small number of 2026 CVEs seeing exploit activity, we could find that by year-end the ratio of old-to-new CVEs seeing exploit activity will be even more pronounced.

Of course, it could be that you've fully remediated all relevant CVEs from previous years; at least those that are resurgent, rather than being exploited for the first time. Still, with the rate IT infrastructure changes for most organisations, it would be prudent to reassess your exposure to older CVEs when new exploit activity is observed.

All of this begs the question: is it really TTE (or even zero days) that we should be worried about? Or should we aim to identify which CVEs (regardless of age) are suddenly popular with cybercriminals, and mitigate them as quickly as possible?

Obviously, this doesn't replace the need for routine vulnerability management work, particularly given the large proportion of current exploit activity that relates to older CVEs. However, it's important to be able to respond quickly to exploit activity.

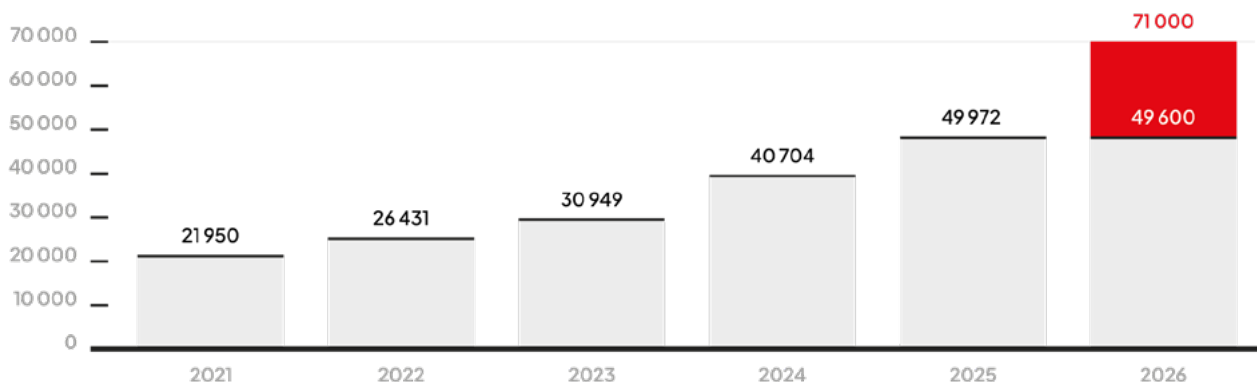
For this, defenders aren't dependent on internet telescopes, but on CISA KEV, EPSS, vendor advisories, and other sources of vulnerability intelligence. Combined, these resources can help defenders identify which CVEs are being exploited right now. And with exposure management, they can also determine if a trending CVE is present and exploitable in their environment, ensuring resources can be allocated to the most dangerous CVEs.

And on that note...

QUESTION 2: HOW IS AI CHANGING CVE EXPLOITATION?

The first, and most obvious, way that AI has affected vulnerability management is through a massive rise in CVE volume. From Anthropic's Project Glasswing and Claude Mythos Preview to a host of other announcements from AI vendors (not to mention the clear correlation between AI model development and pure CVE volume), AI is making its presence felt in CVE discovery.

CVE Volume by Year — CVEs — Predicted



This chart shows what CVE volume by year could look like by EOY 2026 if the current disclosure rate is maintained

How AI is affecting CVE *exploitation* is less clear. The remarkable drop-off in the new-CVE exploit rate that we’ve just noted falls precisely within the period when AI became significant in vulnerability discovery, security research, and ethical and criminal hacking.

We’re firmly into conjecture territory here, so take this with a pinch of salt. We don’t know how many CVEs (of any age) will see significant exploit activity by the end of 2026. Is it possible that, as AI tools are increasingly used by attackers, we’ll see greater clustering of activity around a smaller number of CVEs? If so, that would mirror findings in other AI-affected fields that suggest LLMs are “homogeneously creative”¹⁹ and that AI creative outputs exhibit less “rarity” vs. human outputs.²⁰

Of course, AI-assisted cybercriminals will only rely on specific CVEs for as long as they *work*, so it may be that any response from the security industry will force greater diversity in CVE exploitation. But for the time being, if the question is “are AI-assisted attackers more likely to focus on a smaller range of CVEs compared to human-only attackers?” the answer is a definitive *maybe*.

With that conjecture out of the way, we’ll return to something more evidence-based. The 2026 DBIR attempted to assess the impact of generative AI on cyber threat activity. It found that “most AI-assisted malware development and tooling was associated with well-known and defined attack techniques”, and just “2.5% of AI-assisted malware observations involved less common techniques”.

“AI’s primary impact is currently operational: automating and scaling techniques defenders already know how to detect, not yet unlocking these novel or rare attack surfaces—which means defensive postures don’t need to be reinvented today, but they do need to keep pace with faster, more adaptive execution.”

2026 DBIR

19. [“Large language models are homogeneously creative”](#), E Wenger et al, March 2026

20. [“StoryScope: Investigating idiosyncrasies in AI fiction”](#), J Russell et al, April 2026

This finding closely mirrors what we're seeing in the Bug Bounty world. AI is undoubtedly helping ethical hackers become more efficient and effective, but (at least for now) has very limited capacity for uncovering completely novel techniques.²¹

❗ DOES AI-POWERED = MORE EFFECTIVE?

Any time you're wondering how malicious actors might use AI, it's worth considering the results achieved by legitimate actors in similar fields.

In the Bug Bounty world, it appears that AI tooling adds more value for ethical hackers who are already highly skilled. That makes sense; it's far easier for an expert to assess the quality of AI outputs, discard what doesn't work, and make any necessary changes.

“LLMs are far more useful for an experienced hacker compared to a beginner. An experienced hacker understands what the LLM is really doing, and how to improve it with skills and custom-written tools. But a less skilled hacker doesn't yet have this expertise, so they can't get the same benefits from LLMs.”

Brumens, ethical hacker, developer and Bug Bounty hunter

We're seeing similar phenomena in related fields. In software development, AI tools have variously been found to boost developer productivity by around 10%²², reduce productivity by 19%²³, improve quality metrics by 20–25%²⁴, and also *harm* quality metrics by introducing more code issues²⁵. So it's fair to say the jury is out on how beneficial AI tools are for software development at a macro scale.

21. [“Can AI do novel security research? Meet the HTTP Terminator”](#), Portswigger, August 2026

22. [“How much more productive are AI-powered developers?”](#), GitClear, October 2025

23. [“Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity”](#), METR, July 2025

24. [“AI-Driven Code Review and Quality Assurance: Revolutionizing Software Development”](#), IJRSET, September 2024

25. [“State of the AI vs. Human Code Generation Report”](#), CodeRabbit, 2026

However, anecdotally, there are many accounts of senior developers obtaining significant benefits from AI tooling, particularly for increased productivity. This idea is backed by Google's 2025 DevOps Research and Assessment (DORA) report²⁶, which found that AI "magnifies the strengths of high-performing organizations and the dysfunctions of struggling ones".

What we're getting at here is that, while AI tooling is undoubtedly being used to great effect in cybercrime, it's not a panacea for cybercriminals. Given that legitimate IT disciplines are finding both positive use cases and adverse impacts, it's unlikely that all cybercriminals will suddenly become hacking experts just because they have access to Claude Code.

❗ WHAT EVIDENCE DO WE HAVE OF ATTACKERS USING AI?

Now, that's an interesting question.

IBM's 2026 Cost of a Data Breach report states that AI-driven cyberattacks rose by 56% in the last year, with 25% of organisations that experienced a malicious attack reporting it was AI-driven. Over half (56%) of this activity was deepfake impersonations, while AI-generated malware accounted for 19%. AI-driven attacks also added one million dollars to average breach costs.

This leaves 75% of malicious activity in the dataset either non AI-driven (64%) or undetermined (11%). That's not exactly a slam dunk for the efficacy of AI in cybercrime. Still, attacked organisations wouldn't necessarily be able to detect all AI use, so this may not be a fair representation.

Meanwhile, researchers at the Cambridge Cybercrime Center published a paper in March 2026²⁷ based on their analysis of "large-scale digital sources from the cybercrime underground".

²⁶. ["2025 DORA State of AI-assisted Software Development report"](#), Google Cloud, September 2025

²⁷. ["Stand-Alone Complex or Vibercrime?"](#), J Hughes et al, March 2026

The authors state, “AI is seeing some early adoption in existing large-scale, low-profit passive income schemes and trivial forms of fraud but there is little evidence that it is so far giving rise to widespread disruption in cybercrime.” The report also notes AI is *not* widely used as a “skill multiplier or innovative disruptor” for malicious coding, but is being used for relatively mundane tasks like code pasting and error checking.

And, mirroring our previous chain of thought, the findings suggest that AI tooling is largely beneficial for “already skilled actors, with low-skill actors finding little utility in vibe coding tools compared to pre-made scripts”.

It should be noted that this report focuses on what it calls the “cyber-crime underground,” rather than serious organised crime groups or nation state actors. Given this context, the findings make a lot of sense. The relative value of AI tooling will naturally be reduced for less skilled actors. Still, it’s interesting to note claims of rapidly upskilled junior cybercriminals may have been overblown.

Given our presence in the offensive security space, both from an ethical hacking and AI testing perspective, our view is that skilled actors (including organised crime groups) are undoubtedly benefiting from AI in many of the same ways as skilled ethical hackers. If you’d like to get an idea of what that looks like, our blog contains many detailed articles from top security researchers.²⁸

❌ COST CUTTING FOR CYBERCRIMINALS

Cybercrime is an illegitimate business venture. It aims to make money. And, just like a legitimate venture, it generally requires involved parties to first *invest* money; particularly if they want to scale their operations. As the old saying goes, “you have to spend money to make money”.

A key way that AI affects cybercrime operations is by reducing the cost-barrier to entry for certain attack vectors. Modus operandi that were previously off-limits for small criminal groups (such as reversing a binary to autonomously find weaknesses rather than simply buying an exploit from a dark web marketplace) are increasingly viable options.

²⁸. Like this one: [“How to use Claude Code for Bug Bounty: find fast, validate manually”](#)



These are cost barriers being breached, not just technical barriers. It has always been possible to buy the skills necessary to enact highly sophisticated attacks. It simply wasn't economically viable for most criminal groups. While state-sponsored actors remain more motivated to develop zero day exploits,²⁹ the reduction in TTE suggests new actors may be using AI tooling to adopt methods that were previously beyond their reach.

However, evidence of such "AI upskilling" on a mass scale is currently thin on the ground.

Researchers at Trellix³⁰ have observed an increased commoditisation of AI-assisted attack planning and evasion, as well as a more advanced market for AI tooling. However, Trend Micro's researchers³¹ counter that the outlook for 2026 is "not apocalyptic; it's incremental". They note the criminal AI market is largely focused on jailbreaking legitimate LLMs, and is therefore engaged in a continuous cat-and-mouse game to evade detection and subsequent loss of access.

Since criminal groups don't publish their finances (or pay taxes), there's no way of knowing precisely how much AI tooling may impact their bottom line, or their operational model. Still, it's clear that AI tooling is affecting the economics of cybercrime, and we can expect more clarity on exactly what this means for defenders in the coming months and years.

²⁹. Precisely because they don't have to make money, and can expend far more resources than criminal groups

³⁰. ["Weaponized AI: The Commoditization of Cybercrime"](#), Trellix, August 2026

³¹. ["An Update on the State of Criminal AI"](#), Trend Micro, January 2026

❌ THERE'S MORE TO CYBERCRIME THAN RECONNAISSANCE AND PAYLOADS

Our main concern is not whether AI can be used in cybercrime. It can, and it undoubtedly is. However, it's not enough to wave a hand and say "attackers have AI so they can scale their attacks infinitely in a fraction of the time." Much like the TTE conversation, the reality is more complex.

We've noted that the vast majority of cybercrime is financially motivated. Once a cybercriminal finds a working exploit, they need to monetise it, and this requires scale.

While some aspects of cybercrime lend themselves readily to AI assistance (e.g., reconnaissance, CVE discovery, and payload creation), others don't. There are significant infrastructure requirements for mass cyberattacks that can't easily be handed over to an LLM.

Turning a working exploit into money requires multiple stages, including reconnaissance, weaponisation, delivery, exploitation, installation, command and control, and finally monetisation. A large part of the process is logistical in nature... and logistics are notoriously hard to automate.

In particular, the infrastructure requirements of cybercrime are significant. Botnets, delivery channels, C2, and hosting require upkeep and replacement. AI can potentially be used to make these functions faster and easier, but it doesn't remove the underlying operational requirements or the need for human involvement. And again, how many legitimate organisations have managed to eliminate complex manual work using AI?

Returning to the Cambridge Cybercrime Center report, the authors found almost no evidence of LLMs or automation agents being used to scale botnet management, DDoS, or ransomware logistics. Even if there were, the cybercrime economy is often distributed in nature, with different players fulfilling the various functions required for a complete campaign. This alone would create natural delays in the mass exploitation process.

Again, cybercrime is overwhelmingly profit-focused. Cybercrime groups make risk-based financial decisions just like a legitimate business; they don't invest heavily until they are confident of a return. Given that AI tooling is pay-per-use, it's unlikely they would risk spending heavily without a clear path to profit. There aren't many legitimate businesses rushing to hand over control of their wallets to an autonomous agent, so it's probably not reasonable to expect an organised crime group to do so.

🔪 SUMMING UP

It's not enough to say "threat actors have AI now so everything is instant and infinitely scalable". The process has to be financially viable for the criminal groups involved.

AI is undoubtedly helping attackers discover vulnerabilities more quickly, identify vulnerable targets, and develop exploits. But, at least for now, there's limited evidence that AI tooling is fundamentally altering the exploit timeline for CVEs.



❌ PROTECTING AGAINST CVE EXPLOITATION (AND BIKE THEFT)

Earlier, we asked you to imagine yourself as a 1950s cyclist aiming to protect your bike from thieves armed with a cordless grinder. This poses a challenge, because cordless grinders were newly-invested, so it's highly unlikely you would have been aware of this possibility.

Obviously, you know bike theft is possible, and even that a thief might try to cut your bike chain. But the *specific* threat posed by cordless grinders wasn't foreseeable. Does that leave you defenceless?

A primary function of protective security is to manipulate conditions such that even if an unforeseen (or unlikely) threat arises, the potential for harm is minimised. In this case, you might choose to:

- Avoid locking up your bike in crime-heavy areas
- Keep your bike indoors at night
- Prefer areas with high foot traffic
- Cover your bike when it's left unattended

These are all actions that can significantly reduce the chance of your bike being stolen, regardless of a thief's preferred method. Still, a motivated thief could still steal your bike, and a cordless grinder might render your bike chain close to worthless.

The situation is similar in vulnerability management. You won't always be able to predict which CVEs (new or old) will become popular. And it's clearly not possible to eradicate all CVEs from your IT environment at any given time. So what are your options?

The IT security equivalents of “manipulated conditions” might include:

- General cyber hygiene
- Network segmentation
- Configuration management and enforcement
- Protective security solutions and controls

And naturally, strong detection and response capabilities are essential to prevent harm in the event of an ongoing attack. This should all be a given, but it's important to remember that vulnerability management doesn't exist in isolation.

Still, effective vulnerability management is essential to combat the ever-present threat of CVE exploitation, including when the specific CVEs targeted can't be predicted. Given this, what can defenders learn from the issues raised in this report?

KEY TAKEAWAYS FOR DEFENDERS

The current state of CVE exploitation is more complex than a pure TTE figure suggests. Practically, it doesn't matter if a CVE is a brand new zero day or a forgotten relic of a prior year. What matters to a cybercriminal is whether it's broadly unmitigated. And what matters to defenders is whether it's present and exploitable in their environment.

At the same time, AI is changing CVE exploitation and cybercrime more broadly, though it's not yet clear what this looks like on a mass scale.

No doubt speculation and research will continue unabated, and there may well be more definitive conclusions in the near future. For now, it's worth watching how AI affects related legitimate fields (like ethical hacking, IT infrastructure, and software development) as it's likely that similar benefits will be realised by cybercrime groups.

With all that said, here are five takeaways from this report for IT security teams:

01 SPEED IS IMPORTANT, AND ARGUABLY MORE SO THAN EVER

In the past, it was uncommon for new CVEs to be exploited prior to patch availability or public awareness. Today, it does happen. However, in most cases, you will have some time between CVE publication and significant exploit activity. Vendor advisories and patches generally come quickly, particularly for major threats such as React2Shell.

Simply, the sooner you can assess exposure and implement mitigations once exploit activity is observed, the better. Mitigations deployed within hours or days represent a significant risk reduction compared to mitigations that take weeks or months.

Still, even with consistently rapid mitigations, you could be an unlucky outlier. This is why protective security work beyond vulnerability management, as well as prompt detection and response, are so important.

02 ALTERNATIVE MITIGATIONS MAY BECOME EVEN MORE IMPORTANT

Most large vendors are quick to provide patches for published CVEs, particularly when exploit activity is observed. However, this is not always the case, and it's not ideal to be dependent on a third party for your mitigations.

Patchless alternatives such as firewall, WAF, and IDS rules are already commonly used for CVE mitigation. These benefit from being significantly faster to implement compared to patching. They may prove a more popular first-choice solution for emergency CVE mitigations in the near future, particularly if TTE and unexpected CVE resurgence continue to trend in the wrong direction.

Optimal solutions will vary depending on your IT infrastructure. You

should aim to have a standard process in place for emergency CVE mitigation that does not rely on vendor patch availability.

03 PROACTIVE IS BETTER THAN REACTIVE

It's always better to pre-empt a threat rather than react to it. In addition to the routine work of vulnerability management, it's a good idea to pre-emptively mitigate CVEs that are *likely* to be exploited in the near future.

This is where resources such as EPSS and vulnerability intelligence come in. Using a variety of sources, it is possible to identify the CVEs most likely to be exploited in the near future, enabling your organisation to pre-emptively identify dangerous and relevant CVEs to mitigate in your environment.

If you can avoid the need to rush through mitigations in response to observed exploit activity, that will naturally be preferable from a risk management perspective.

04 SPOTTING EARLY WARNING SIGNS IS CRUCIAL

It's not always possible to be proactive, so identifying which CVEs are being exploited *right now* is a big deal. You can't reasonably expect to always predict which new CVEs will prove popular, or which old CVEs will be resurgent.

Thankfully, there are structural limits to how many different CVEs can be mass-exploited at once, and cybercriminals are financially incentivised to focus on a small number of highly effective CVEs at any one time.

It is both crucial and realistic to identify those CVEs promptly once exploit activity is observed. There are various ways to achieve this, but whatever you do, you should have this capability, and it should feed directly and promptly into your emergency CVE mitigation process. Again, in these emergency cases, we're aiming for mitigation within hours or days, not weeks.

05 UNDERSTAND WHAT MAKES A CVE DANGEROUS FOR YOU

The most dangerous CVEs have three things in common. They are:

- Present in your IT environment
- Externally exploitable in context
- Currently being (or likely to be) exploited in the wild

CVEs that are not present or not exploitable in your environment do not pose a risk to your organisation, even if they are being widely exploited. This is where vulnerability mitigation meets **exposure management**, and requires a thorough, current understanding of your attack surface.

CLOSING THOUGHTS

The findings and recommendations in this report are based on data from 2025 and H1 2026. Given the rapid rate of AI advancement, it is possible that things will change in the coming months.

With that said, it's important not to be swept away by unnecessary scare tactics or unproven claims. While it's far from the full picture, TTE is falling, and cybercriminals are using AI in various ways to improve their speed and efficiency. However, at least for now, there is no substantial evidence that mass exploitation is likely prior to CVE publication.

It is, of course, important to continue monitoring CVE exploit activity to see if it really does become unmanageable using current processes. For the time being, sensible pre-emptive controls and prompt response to CVE exploitation are both crucial and valid as a strategy for mitigating CVE risk.

UNDERSTAND & MITIGATE CVE EXPOSURE WITH YESWEHACK

In this paper we've advocated for an emergency process for prompt mitigation of high-risk CVEs. For this to be possible, you need to be able to determine whether a CVE is:

- Present in your IT environment
- Externally exploitable in context
- Currently being (or likely to be) exploited in the wild

Answering these questions requires several capabilities: a real-time understanding of your attack surface, vulnerability intelligence that can identify the most dangerous CVEs, and the ability to validate exploitability in the context of your IT environment.

FOR PRE-EMPTIVE AND EMERGENCY CVE MITIGATION

Exposure Management helps organisations quickly identify dangerous CVEs and assess their exposure. It continuously maps the full attack surface, identifies potential exposures, and validates them using fully documented attack scenarios, called Checkpoints.

It works on three principles:

1. **Continuous discovery.** The platform continuously maps your assets and fingerprints their components across the entire attack surface.
2. **Validation.** Each Checkpoint is a fully documented attack scenario, validated by our experts. So you only receive alerts when there's a genuine, reproducible exploit path.
3. **Speed.** Checkpoints focus on vulnerabilities (including CVEs) that are currently being exploited or likely to be exploited imminently. They are developed based on our years of vulnerability intelligence, cross-platform Bug Bounty insights, and external sources.

With Exposure Management, your team can:

- Rapidly assess CVE exposure, and prioritise dangerous CVEs for immediate mitigation.
- Respond quickly and confidently to new, high-profile CVEs like React2Shell and wp2shell.
- Focus ONLY on CVEs that are present and exploitable in your environment, with **zero noise**.

When a new high-impact vulnerability appears, we immediately build a Checkpoint that reproduces the full attack scenario, deploy it, and test your full attack surface. If you're vulnerable, it's detected within minutes and you're notified automatically.

Find out how Exposure Management could help your organisation protect against high-risk CVEs:

BOOK A DEMO →