TRAVELERS

Q3 CYBER THREAT REPORT

A Coordinated Campaign: Ransomware Rises Again

2025

Table of Contents

Published by Travelers with contributions from:

Ryan Bell
Director, Threat Intelligence – Cyber Risk Services, *Travelers*

John Lippe
Director, Cyber Claim Forensics, *Travelers*

Josh Doguet
Sr. Manager, Incident Response Intelligence, *Travelers*

Nicholas Kelley-Ossey
Sr. Director, Cybersecurity, *Travelers*

Aleesha Quintana
Cybersecurity Technologist, *Travelers*

Alex Pinto
Sr. Director of Product Marketing – Cyber, *Travelers*

Executive Summary

The ransomware ecosystem took on a new form in the third quarter. Earlier this year, it had fallen into disarray after a prominent group, RansomHub, dissolved thanks to the arrests of key members. That led to a quarter-over-quarter decline in ransomware activity in Q2, the first we'd seen in more than a year. The question was whether that disruption would prove durable – and, in the likely case activity did return, what form it would take.

Now we have the beginning of an answer. As we cover in detail in this quarter's report, a campaign of attacks on virtual private network (VPN) software took place that had some unusual characteristics. Rather than focusing on a newly discovered vulnerability, the campaign attacked weaknesses that were already well known. And the suddenness and persistence of the campaign gave the impression of strong coordination, rather than the haphazard patterns of activity we typically see when a new idea ripples through the cybercrime world.

This new complexion of the ransomware ecosystem is paired with risks arising within the tech stacks of many organizations. In this edition we also look at how the rise of agentic artificial intelligence (AI) tools could expose organizations to what's been called the "lethal trifecta" of AI risk. Unlike our guidance for defending against the VPN attacks we saw in Q3, defending employee AI usage isn't about following through on existing best practices – it may require a new way of thinking about security.

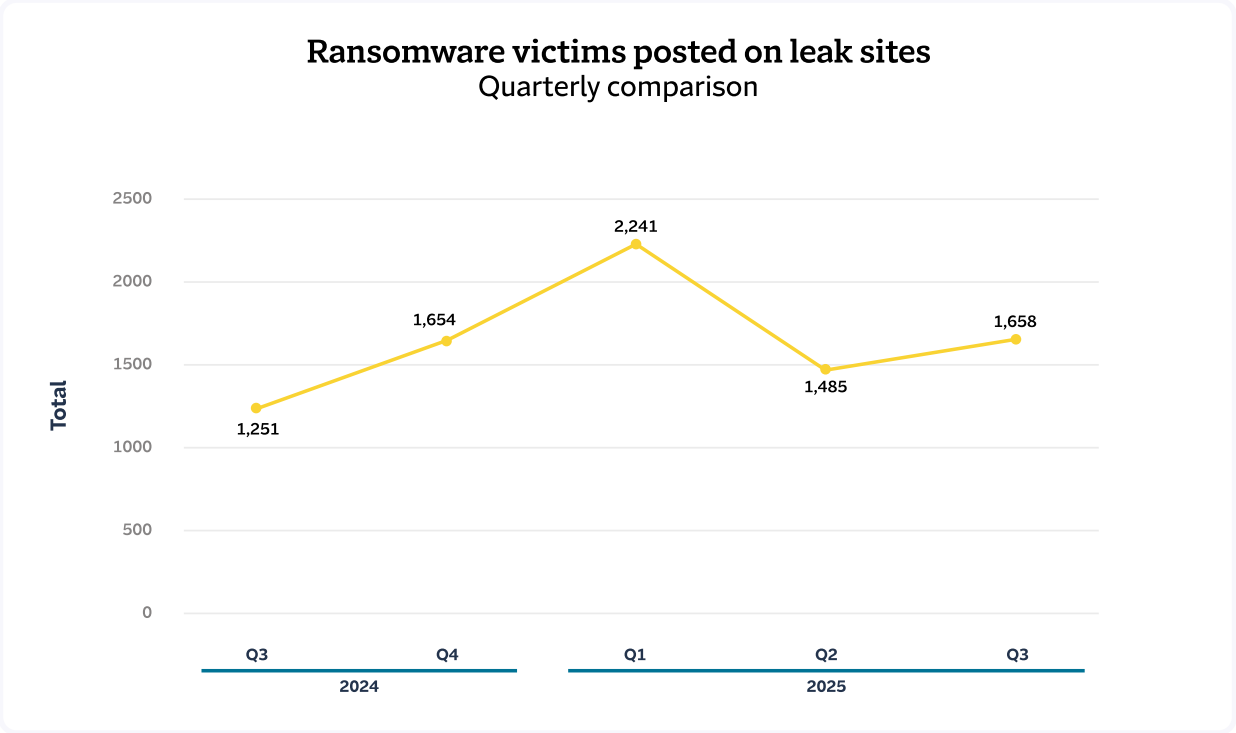
Ransomware activity returns to trend: 1,658 victims were posted to leak sites in Q3, making it the second-highest quarter by this measure going back to the beginning of our data set in 2021.

A campaign of attacks hit users of a popular VPN provider: the Akira ransomware group and its affiliates were responsible for more than 50% of all Travelers ransomware claims in August and September.

AI usage is creating a potential new arena for attacks: Travelers data indicates sanctioned AI adoption is growing rapidly; with new agentic AI tools, organizations must consider a different set of risks.

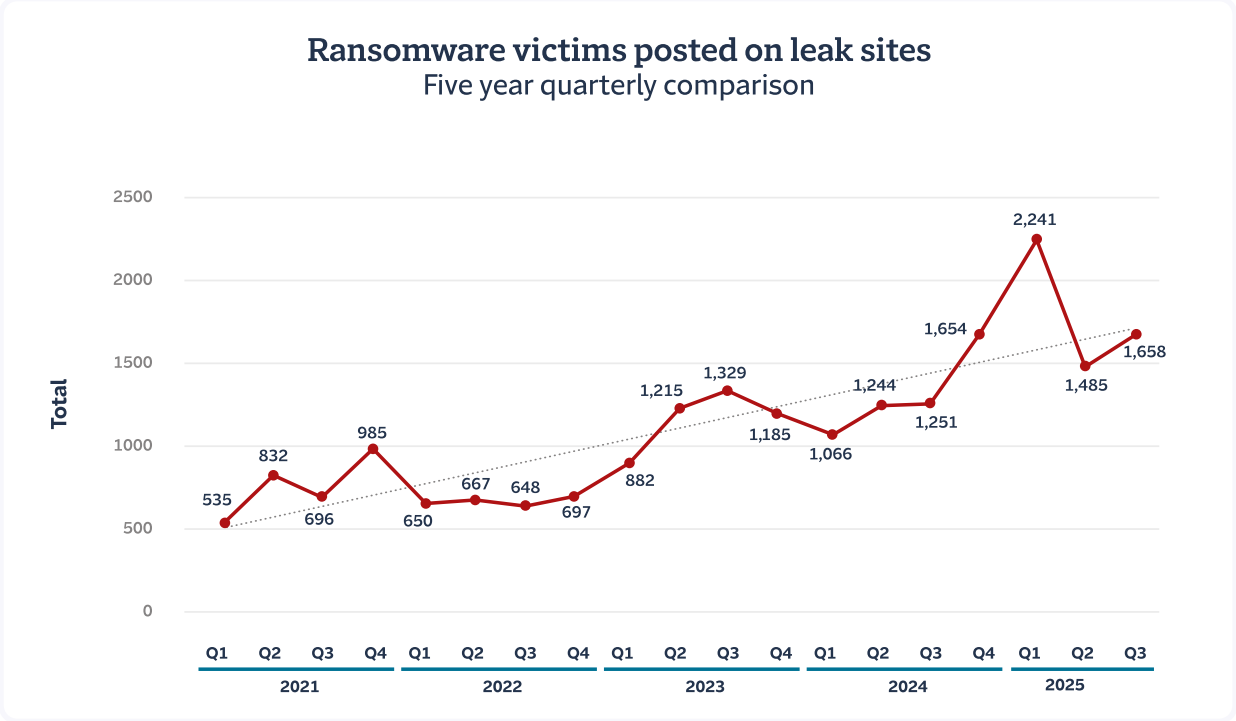
Ransomware Leak Site Activity Rises Again

The reprieve didn't last long: after we saw a quarter-over-quarter dip in Q2, ransomware activity* rose again in Q3 to reach the second-highest level in our database going back to 2021. The quarter's total of 1,658 was second only to the spike in activity we saw in the first quarter of this year.



*Data shared on leaksites provides a proxy for overall ransomware activity. A victim's information will typically be posted if the victim has refused to pay a ransom. This means that the data should be viewed as a fraction of overall activity, but one that can provide a longitudinal comparison of activity over longer time frames.

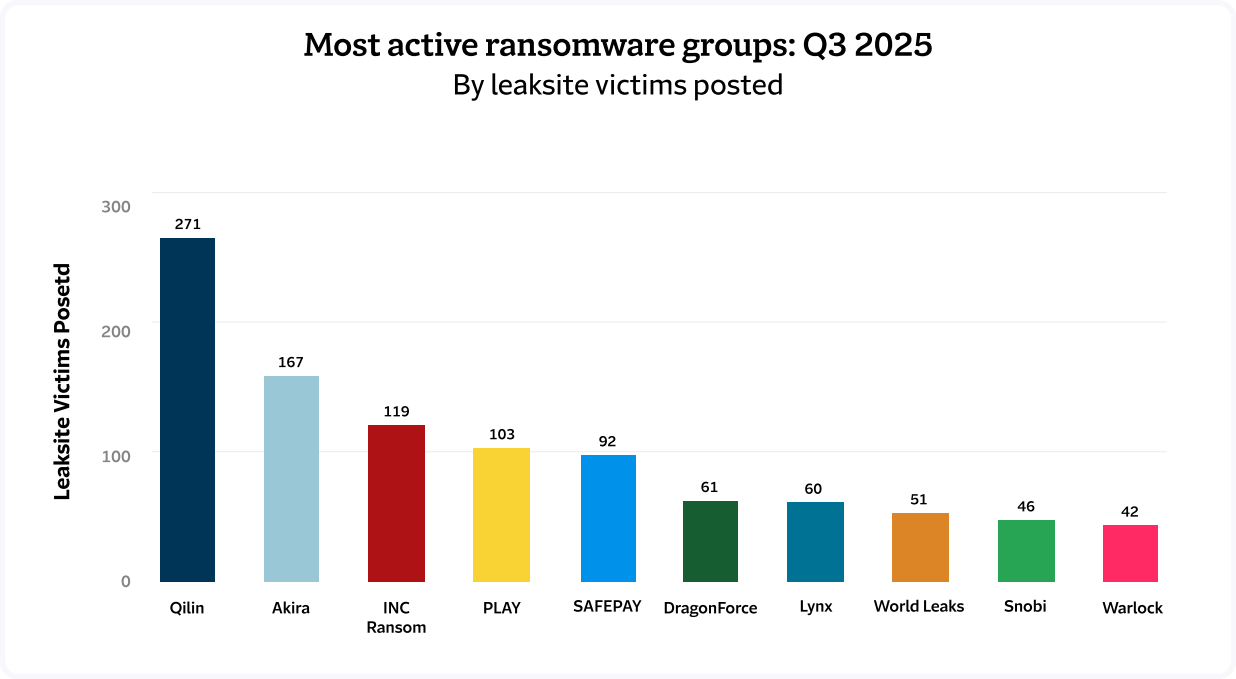
As we noted in our last report, we viewed the second quarter's drop in activity as likely a short-term response to changing conditions in the ransomware ecosystem, not a departure from the long-term trajectory. After seeing another quarter's data, it looks like our bearish view has been upheld. The third quarter sits right on the long-term trendline, and our view still holds that unless there is a major change in the law enforcement posture within the key countries where most cyber criminals are based — or a drop-off in the monetary returns from these attacks — ransomware activity will continue to rise over time.



Contributing Factors to Activity in Q3 2025

Throughout Q3, we observed activity from **eighty distinct ransomware groups**. That's the most simultaneously active groups we've seen in a quarter. Yet despite the proliferation of new groups, we are still seeing the dynamic of a handful of groups making up a large share of the activity in each quarter. The two most prolific groups in Q3 — Qilin and Akira — together conducted about a quarter of the ransomware attacks, while the top 5 most active groups accounted for 45 percent. Qilin claimed 271 victim organizations on its dark web leak site in Q3, while Akira listed 167 victims throughout the quarter.

After RansomHub's sharp decline in early April, Qilin significantly strengthened its position by recruiting many former RansomHub partners and participants, offering them attractive conditions and opportunities. Qilin claimed over 15 percent of all attacks from April to September, while Akira's share was around 10 percent.



Akira's Campaign

At the end of July 2025, security researchers observed sustained and increasing activity by Akira targeting firewalls via Secure Sockets Layer (SSL) VPN login requests, with malicious logins followed within minutes by port scans and rapid delivery of ransomware.

This campaign by Akira targeting VPN accounts intensified through Q3, with new related infrastructure observed as recently as September 20, 2025. The attackers managed to successfully authenticate themselves on accounts with multifactor authentication (MFA) enabled for one-time passwords. See the next section for more on this campaign.

VPNs Were Under Attack in Q3

In early August, our team began noticing an uptick in ransomware attacks targeting customers of one popular VPN provider. This was noteworthy, but not unusual: we have seen clusters of attacks targeting specific remote access technologies before.

But the attacks kept coming. By the end of September, **more than 50% of Travelers' ransomware claims** during this two-month period had been attributed to attacks exploiting vulnerabilities in this single technology, including more than 70% of our ransomware claims in August. Early on, nearly all of the attacks appeared to be executed by the Akira ransomware group and its affiliates, though later in the quarter we did see other groups beginning to exploit the technology in similar fashion. Over that time our team issued targeted threat alerts to nearly 2,000 policyholders we identified as users of the technology using our risk scan.

This was an unusual level of concentration on one technology or attack pattern for a two-month span. The only past events that compare involved a zero-day vulnerability — this one did not — and in those cases, more than one group of threat actors was typically involved. We now consider this to be a coordinated campaign by Akira, which represents a departure from typical patterns of ransomware activity.

> By the end of September, more than 50% of Travelers' ransomware claims during this two-month period had been attributed to attacks exploiting vulnerabilities in this single technology

A Soft Market (for Affiliates)

Akira operates as a ransomware-as-a-service (RaaS) operation, meaning the core group develops the ransomware tools and infrastructure while "affiliates" — independent threat actors who pay for access to the ransomware and supporting services — carry out the attacks. What made this campaign especially potent was what appears to have been an aggressive affiliate recruitment push in the months leading up to the attack surge. Akira's expanded affiliate network was likely bolstered by the presence of experienced operators from recently-disrupted groups like RansomHub, which, as we covered in the last edition of this report, was broken up after law enforcement actions earlier this year.

In addition to having numerous affiliate groups available to work with, it appears that there was centralized distribution of the attack playbook and some level of coordination on timing. Except for zero-day situations, we typically see a more organic-looking attack pattern as ideas spread by digital word-of-mouth, with various groups of threat actors sporadically targeting different types of victims over a period of months or years. This was a sudden outburst of activity by comparison.

Inside the Attacks

The attacks exploited multiple points of weakness in the VPN platform's SSL VPN functionality and the way it was set up at various victim organizations.

As we mentioned, it appears that threat actors had been sitting on knowledge of these vulnerabilities for an extended period, conducting reconnaissance and preparing for a coordinated strike. One known vulnerability that is likely to have been involved in the attacks was first disclosed in August 2024, nearly a year before the concentrated attack activity began.

A key factor that likely amplified the impact of the attacks was a configuration oversight following system upgrades. According to security researchers, many organizations apparently failed to reset local administrative passwords after upgrading their VPN infrastructure, leaving default credentials in place. This would have created an easy pathway for threat actors to gain access through a simple brute-force attack, even if security patches had been applied.

When default or otherwise simple passwords are the only line of defense, without a phishing-resistant multifactor authentication process in place, the attacks do not require specialized skill to carry out. That further widens the pool of potential affiliates to work with. It's also possible that some attacks stemmed from previous deployments of malware that harvested credentials from the VPN technology, which could have led to many valid credentials circulating among the threat actors, obviating the need for a brute-force attack.

Unlike ransomware deployments that focus primarily on encryption, some Akira affiliates systematically targeted and destroyed backup systems before deploying their primary ransomware attack. This "backup destruction" protocol significantly reduced victims' recovery options and increased the likelihood of ransom payment. It appears that the effectiveness of this approach was documented and distributed among affiliates, creating a standardized methodology that amplified success rates across the network.

Lock Those Doors

What makes this campaign particularly chastening is the possibility that many attacks stemmed from fundamental security oversights rather than sophisticated hacking exploits. The failure to reset default administrative passwords after system upgrades, the absence of multifactor authentication on critical infrastructure and the lack of proper credential management created opportunities that did not require great technical sophistication to exploit.

Despite being the "front door" to corporate networks, VPNs and other remote access points are under-secured or simply forgotten about with surprising regularity. Many organizations take a "set and forget" approach, applying patches when required but overlooking the basic authentication controls. These are easy targets to find: being exposed to the public internet means they are visible to off-the-shelf scanning technologies. And deploying brute-force or credential stuffing attacks against these targets is relatively low-cost, so threat actors can take many shots.

The Akira campaign succeeded not because of unheard-of hacking techniques, but because of the discipline to systematically locate many targets with similar profiles.

The Akira campaign succeeded not because of unheard-of hacking techniques, but because of the discipline to systematically locate many targets with similar profiles. Basic hygiene like enforcing strong passwords and implementing multifactor authentication could have stopped many of the attacks.

This is the most organized effort to date in the trend toward threat actors using more repeatable tactics. The fact that the trend has only gained adherents over time highlights a broader challenge in cybersecurity: the gap between knowing what should be done and ensuring it gets done consistently. The technical solutions to prevent VPN compromise are well understood. The operational challenge lies in maintaining security discipline in infrastructure management, especially during system upgrades, and in day-to-day administration when security considerations can be overshadowed by the urgency of responding to business needs.

Briefing: The "Lethal Trifecta" of Artificial Intelligence Tools

In a recent quarterly report, we described how AI is beginning to amplify the ransomware ecosystem, with threat actors automating what were once labor-intensive steps in the attack chain. They are using AI tools to develop grammatically correct phishing email content in the target recipient's language, and spoofing voices and likenesses for "deepfake" social engineering on phone and video calls.

But that's not the only angle by which to consider AI's impacts on cyber risk and security. There's also the **fact that legitimate, intended AI usage by employees within organizations is expanding the attack surface**, and the vulnerabilities that arise from this usage are only just being discovered.

Adoption of general AI tools is happening more rapidly than businesses can understand the risks. Data from the Travelers Cyber Risk Scan showed a **60% increase in adoption of AI over a two-month span this year** across a large sample of organizations — and that only examined one AI company's products. It also did not account for unofficial "shadow IT" usage by employees.

The data suggests a pace of adoption of AI tools that's arguably happening faster than any shift we're seeing in threat actors' attack patterns due to AI.

AI adoption brings new risks...

As employees explore creative uses of AI, they are also potentially allowing the tools to access sensitive data. And they are potentially using tools with no limitations on what instructions they will follow, and that — as with buzzy "agentic" tools — can carry out actions automatically.

Those aren't separate, independent risks. In fact, those three factors have been described as the **"lethal trifecta" of AI systems** by software developer and researcher Simon Willison. According to Willison, **access to private data, exposure to untrusted instructions and the ability to act or share information externally are a uniquely perilous combination**. When all are present, a threat actor could have significant leverage. But if organizations can clamp down any one of the three, the risk of exfiltration or attack falls dramatically.

What sort of format would an attack targeting this trifecta follow? The most cited concept is **prompt injection**, another term Willison helped to popularize. It involves a threat actor sneaking a prompt into an AI model they shouldn't have access to, enabling them to use it for their own ends.

Prompt Injection

Setting aside the technical-sounding name, the idea behind prompt injection is easy to grasp: a threat actor finds a sneaky way to tell an AI model to do something. They could sneak malicious instructions into the footnotes of a lengthy PDF report, or hide them in an image in a way that's invisible, but machine-readable. The instructions could be to search a database for a certain type of file and send it to an email address controlled by the threat actor, for instance. If an unsuspecting, legitimate AI user uploads the contaminated media to their AI tool, the actions could be executed.

Such attacks have been demonstrated repeatedly by security researchers. At the 2025 Black Hat conference, for instance, researchers demonstrated an exploit of ChatGPT's integration with Google Drive by sharing a file containing hidden instructions with a targeted user. When the victim instructed ChatGPT to process the file, the embedded instructions were executed, allowing the researchers to search the victim's Google Drive for sensitive data and steal it. Researchers have uncovered many similar "vulnerabilities" across the growing ecosystem where large language model (LLM) applications are interfacing with traditional web applications that store data.

Thankfully, to date most examples of prompt injection have come from security researchers rather than from victims of a crime. But it's not the only way that AI can become part of an attack. Our team at Travelers recently encountered a situation in which **a common AI-based productivity tool was leveraged in a social engineering attack to enable the exfiltration of data**. From the outset, the attack was a typical example of a business account compromise, accomplished through phishing; the twist was that the compromised account included access to the AI tool, which the threat actor used to efficiently locate the most valuable files to exfiltrate once inside.

...And a new defense paradigm

Avoiding the lethal trifecta involves regulating a technology that behaves inconsistently even when it's working correctly. There are a dizzying number of variables inherent in the behavior of new models being released, in the different training corpuses that might be used at different organizations, and in the ways the technology might choose to interact with other applications. That means governance needs to account not just for who is allowed to use the tool — as you might control access to traditional IT — but also what that user is allowed to give the tool, and how it is allowed to respond.

For most organizations, reducing risk starts with simplifying the tool's role. They should limit its access to data. They should prevent it from acting autonomously — especially from sending messages, executing code or updating records without human review. And they should set clear boundaries for employees about what information can be sent into public or third-party AI tools.

Aside from governance, other important areas to consider include:

- Monitoring for inappropriate use of Model Context Protocol (MCP) — these are increasingly popular tools that automatically route a request to one of several AI models, and can introduce even more unpredictability into a system that is already difficult to constrain.
- Stress-testing models against prompt injection attacks through "red teaming" — even if the AI solution has been tested by its creators, each implementation may introduce unforeseen gaps that should be tested.
- AI Model Security for detecting and preventing unsafe code execution — fighting fire with fire by leveraging AI to monitor certain types of requests.

The frontier is expanding

Prompt injections have been a focus of security researchers since practically the moment ChatGPT was unleashed to the world in 2022. Why the concern now?

It's only recently that the promise of "agentic AI" is becoming a reality at a mass scale. Now that AI is being built into solutions like web browsers or file management systems, the connection of the least-common leg of the Lethal Trifecta, "ability to act externally," is becoming more widespread, making de-risking the other two legs suddenly much more important.

Conclusion

The third quarter of 2025 marked a point of evolution in cyber crime, with ransomware activity rebounding and taking on a more coordinated, strategic character exemplified by Akira's campaign of attacks. This shift toward organized, sustained attacks on well-known weaknesses, combined with the rapid adoption of agentic AI tools that introduce the "lethal trifecta" of data access, instruction exposure and autonomous action capabilities, signals that organizations must now defend against both more disciplined threat actors and entirely new attack vectors.

Recommendations from the Travelers Cyber Risk Services Team

To mitigate these risks, organizations should adopt a strong cyber prevention program, including the following recommendations detailing the top security investments with the greatest return on investment.

These recommendations will help increase the bar required for ransomware actors to successfully carry out an attack on an organization.

They include:

- Implement phishing-resistant MFA for all remote access and email.
- Run an effective vulnerability management program to quickly patch critical vulnerabilities in edge devices, such as virtual private networks (VPNs).
- Ensure you have reliable backups and have a resilient disaster recovery and business continuity plan.
- Run endpoint detection and response (EDR) solutions with 24x7 active monitoring.

travelers.com

One Tower Square
Hartford, CT 06183

Travelers analysis was made possible with supporting data from eCrime.ch.