# Cloud Threats Retrospective 2026

How systemic weaknesses and AI amplified the impact of proven cloud threats

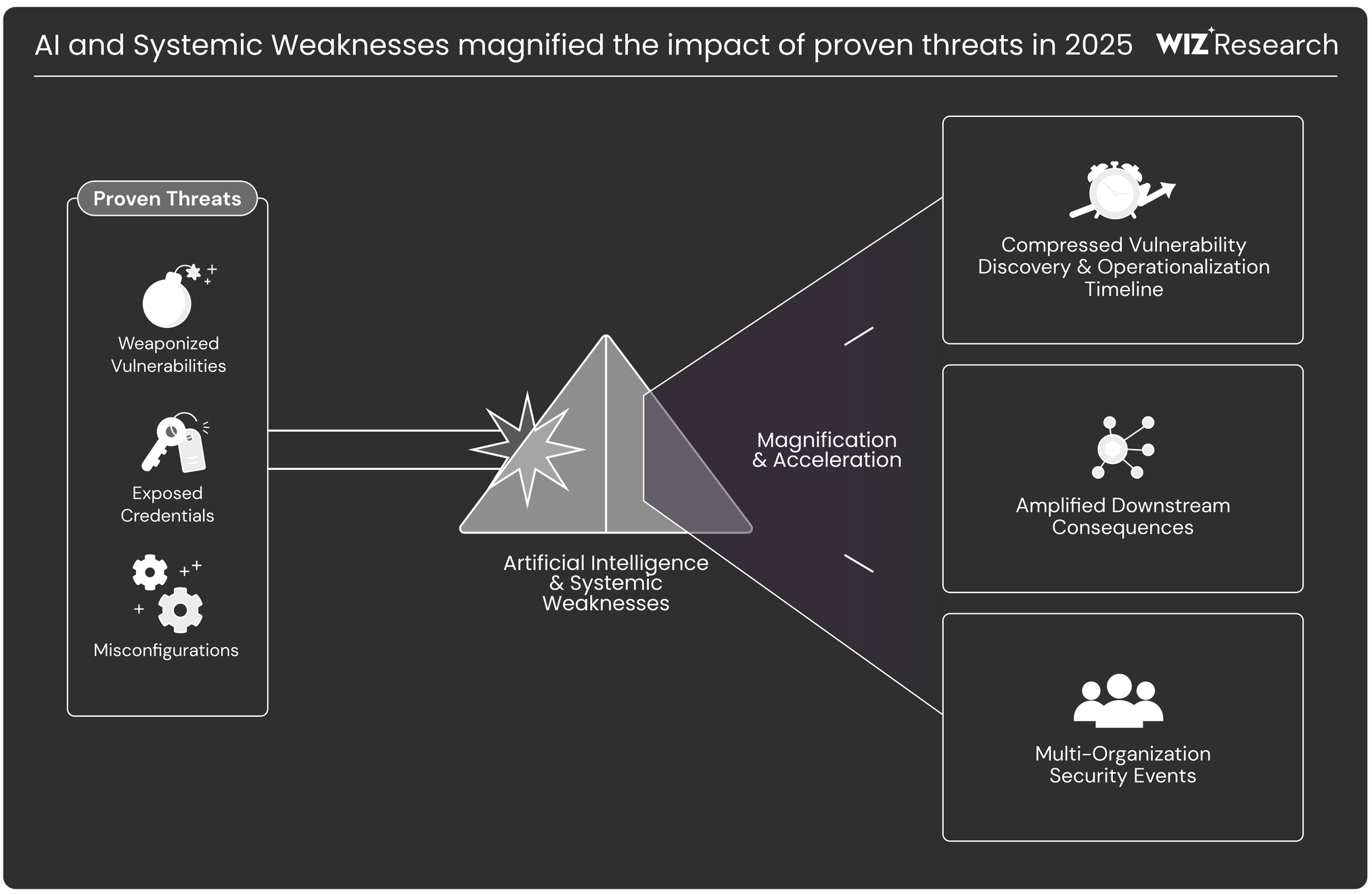
# Table of Contents

# Introduction

Classic weaknesses remained the dominant factor in cloud threat actor activity in 2025. Across publicly documented incidents included in Wiz's Cloud Threat Landscape, initial access most often involved weaponized vulnerabilities, exposed secrets, and misconfigurations.

But the familiarity of these vectors should not be mistaken for limited impact or stagnation in attacker behavior. Beyond individual intrusions, critical incidents such as Shai-Hulud and React2Shell demonstrated how systemic weaknesses across shared infrastructure, software dependencies, and trusted integrations could be weaponized to produce outsized impact. These events showed how inherited trust and ecosystem-wide exposure can amplify the consequences of otherwise well-understood attack techniques.

AI also influenced cloud-focused threat actor activity in 2025, not by introducing fundamentally new attack techniques, but by expanding cloud attack surfaces and enabling threat actor workflows in select cases. **As AI-driven infrastructure, tooling, and automation became more common, familiar security risks increasingly appeared in new contexts and at greater scale.**



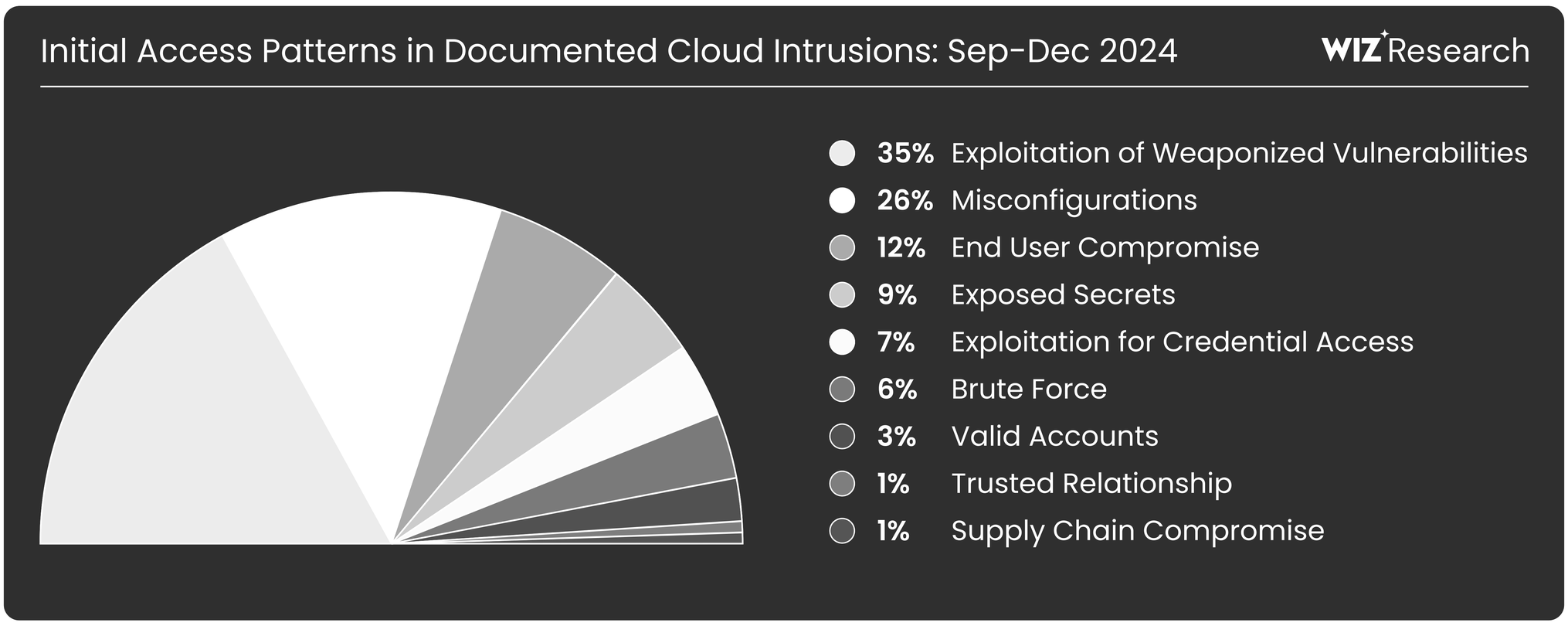

In this report, we analyze threat actor behavior data and real-world incidents to explore how proven cloud weaknesses, systemic trust relationships, and AI-driven change shaped threat actor activity in 2025.

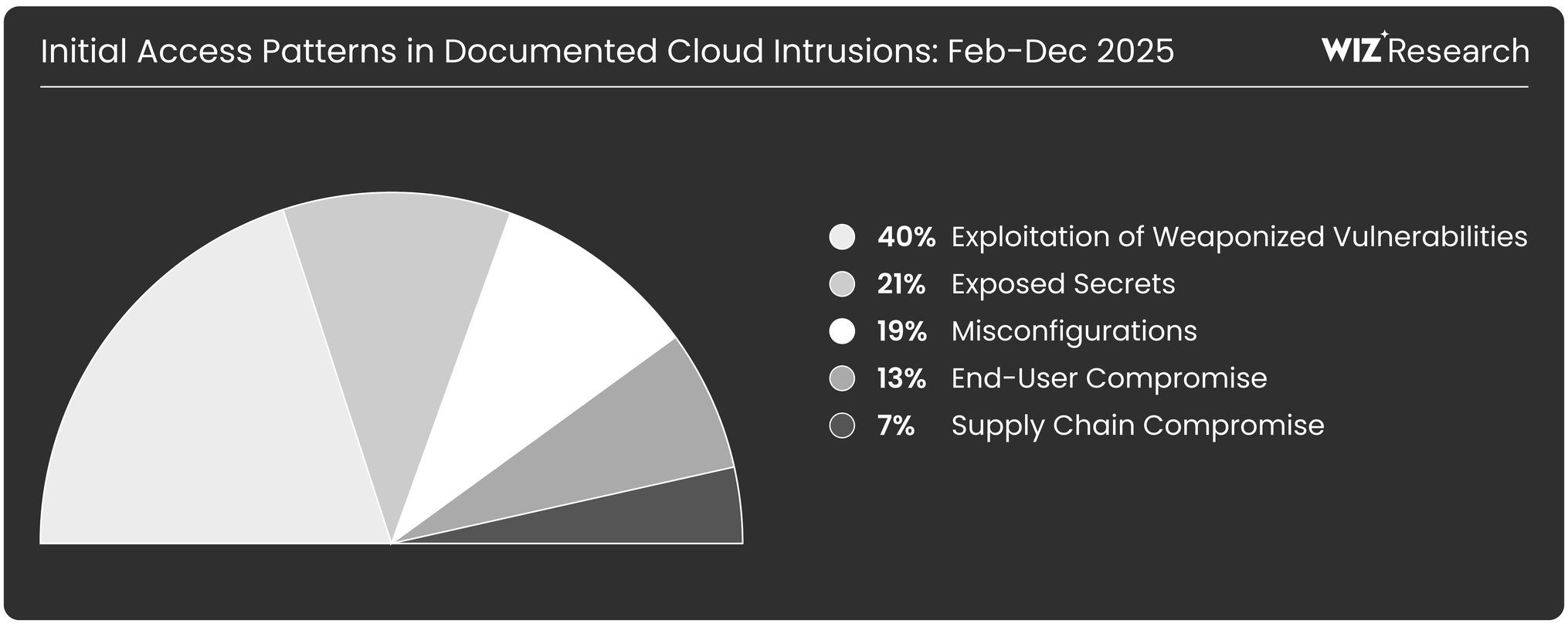
# Threat Actors Continued to Prioritize Proven Cloud Attack Methods

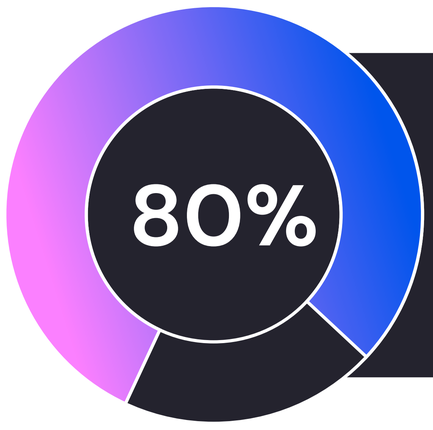
## 1 Analyzing Publicly Documented Intrusions

To understand how cloud intrusions typically occur, Wiz Research analyzed cloud security incidents from 2025, drawing on publicly documented cases in the Wiz Cloud Threat Landscape.

Across this incident set, threat actors continued to rely on a small number of well-known vectors to gain initial access, largely consistent with patterns observed in the prior reporting period. The most common entry points were not novel cloud-specific exploits or advanced identity bypass techniques, but familiar weaknesses in exposure management, credential handling, configuration, and end-user security.






**of documented cloud intrusions in 2025 began with vulnerabilities, exposed secrets, or misconfigurations.**

Across periods, vulnerabilities, exposed secrets, misconfigurations, and end-user compromise persisted as the most observed entry points. While misconfigurations saw a decline in prominence, exposed secrets more than doubled to 21 percent and supply chain compromises rose to 7 percent.

## 2 Where Attackers Are Most Often Detected in Cloud Environments

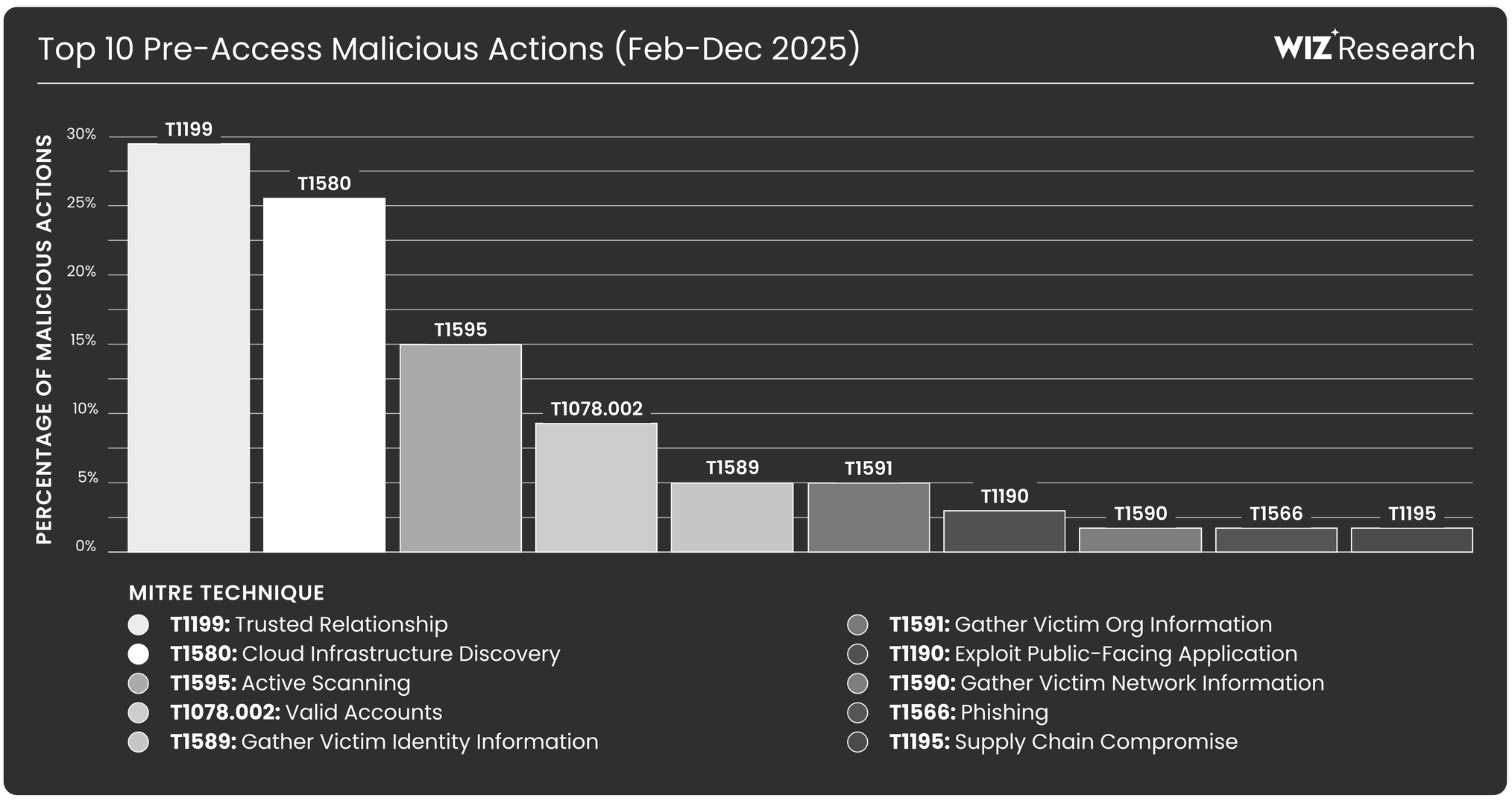
In addition to publicly documented intrusions, Wiz Research analyzed malicious activity observed in Wiz-monitored cloud environments to understand early-stage attacker behavior.

The data below reflects **high- and critical-severity malicious actions** detected by Wiz. These severity levels are assigned to behaviors that indicate elevated risk, such as interaction with cloud control planes, identity systems, or sensitive resources. Detection of these actions does not always indicate successful compromise, persistence, or operational impact.

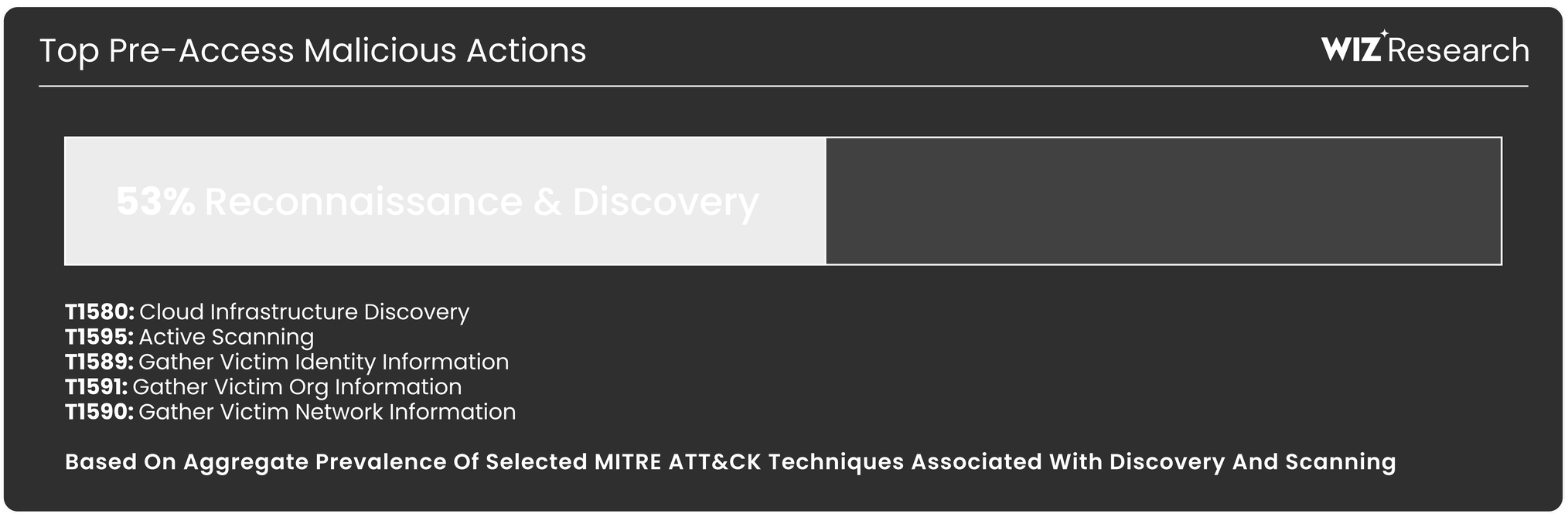
Percentages represent the relative prevalence of each malicious action type per tenant per month.

### Pre-Access Activity

The following chart summarizes high- and critical-severity malicious actions most often detected during the early phases of an operation.



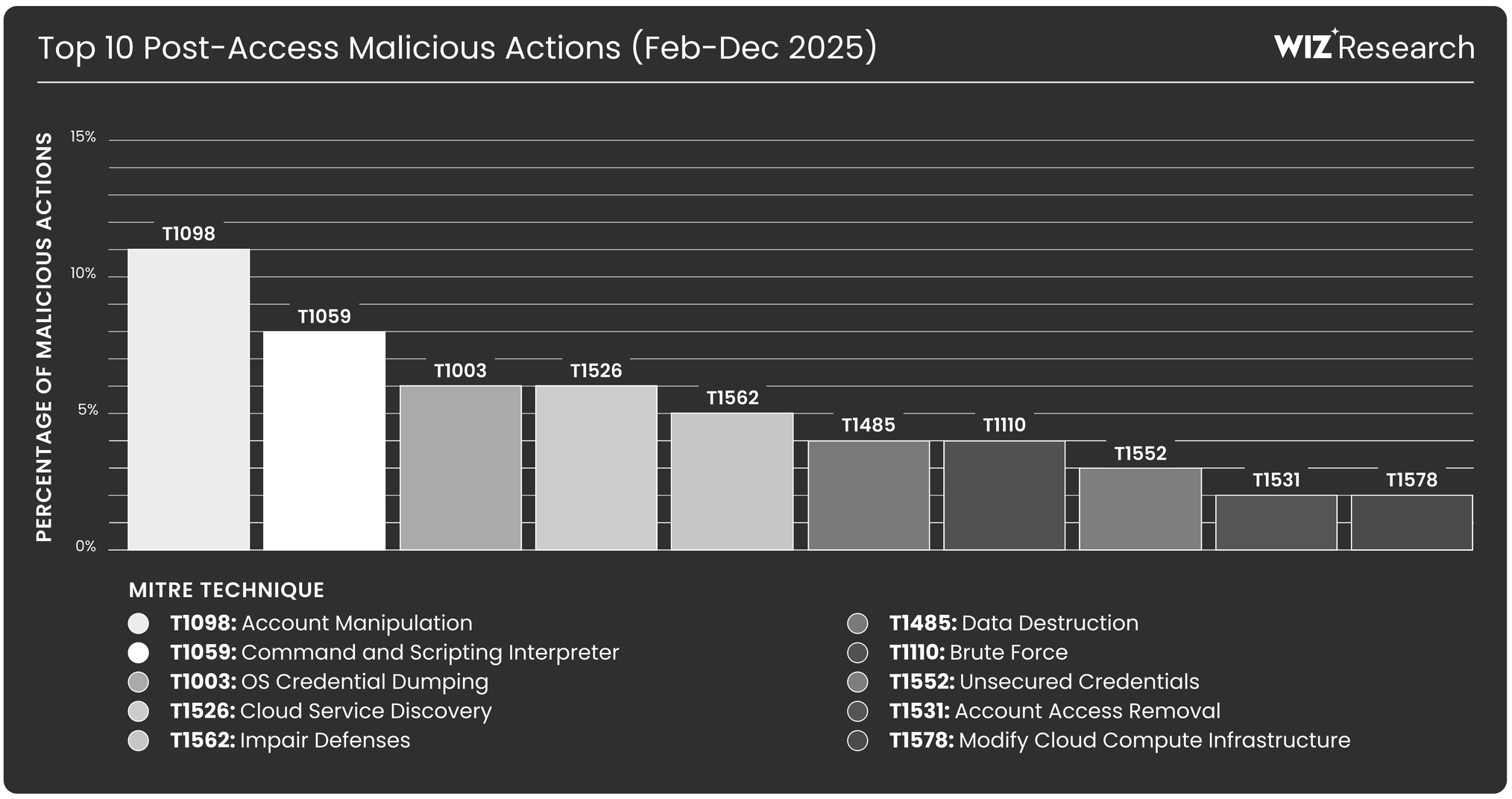
The early part of the operations require malicious actors to not only gain some level of privileged access to a network, but conduct internal reconnaissance to understand where they are and how to accomplish their goal. This creates an opportunity for defenders to identify malicious activity before they are able to accomplish their goals.




Reconnaissance and discovery-related techniques account for over half of observed early stage detections, highlighting how heavily threat actors invest in mapping environments and testing trust boundaries and giving defenders a great opportunity.

## Post-Access Activity

When some level of access is obtained, high- and critical-severity detections cluster around a consistent set of cloud-native behaviors. The chart below highlights post-access malicious actions detected in cloud environments, including short-lived or contained activity.




These actions reflect how threat actors attempt to interact with cloud identities, control planes, and services once access is obtained. Rather than relying on traditional host-based lateral movement, threat actors prioritize identity manipulation, service enumeration, scripting, and resource control, making these behaviors critical early indicators of escalation attempts.

## 3 What These Patterns Reveal

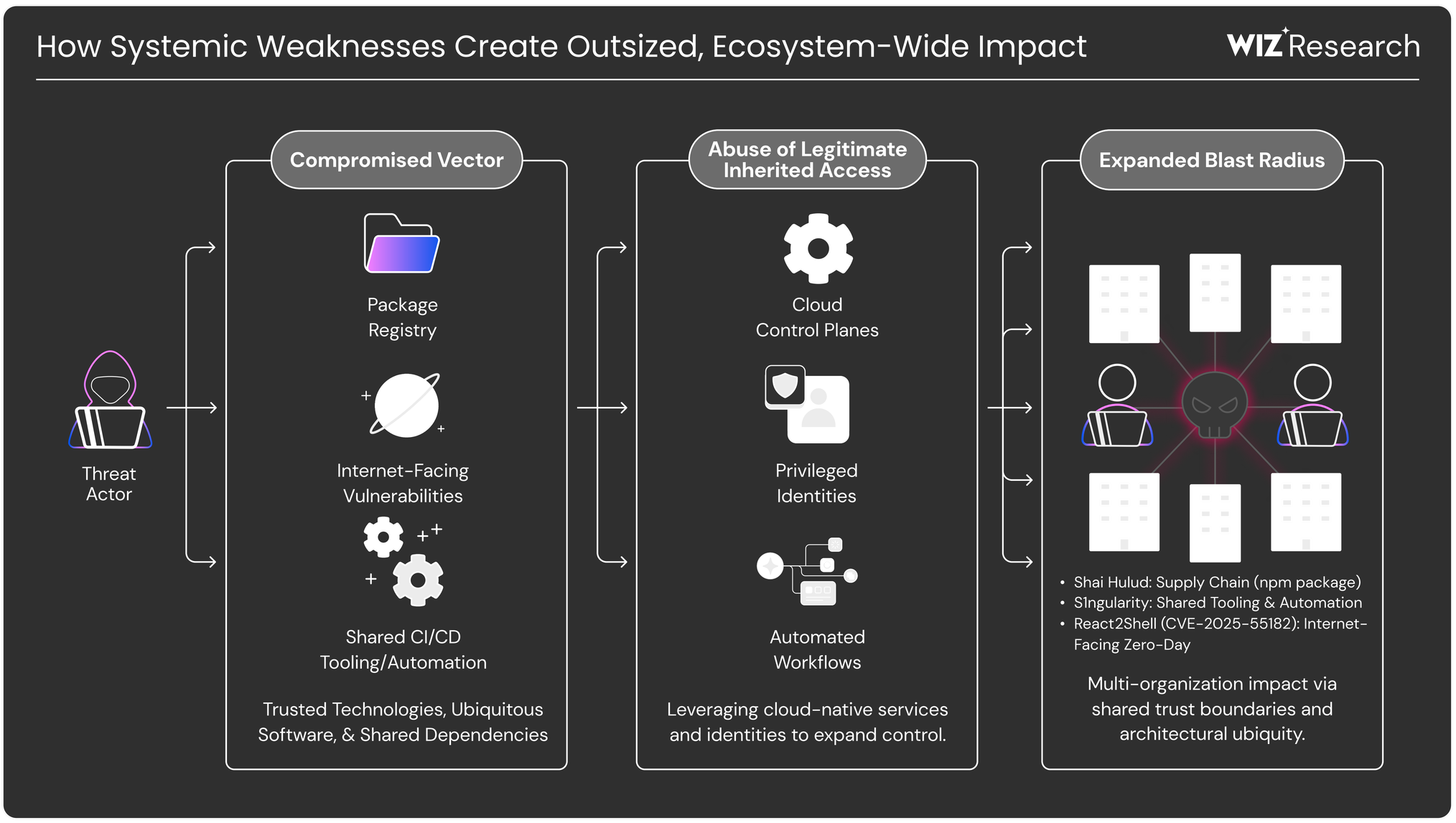
Public breach analysis and detected malicious activity point to some consistent patterns in threat actor behavior targeting cloud environments during 2025.

Publicly documented breaches show that initial access most often stems from vulnerabilities, exposed secrets, and misconfigurations. Within cloud environments, the high- and critical-severity malicious actions most frequently detected reflect the first steps threat actors take to prepare for access, test trust boundaries, and attempt to expand control.

At a high level, these behaviors are not novel. But as we will discuss throughout this report, they continue to carry disproportionate risk when they intersect with exposed services, identities, and trusted relationships. In cloud environments, these actions are often detected early in the intrusion lifecycle, enabling security teams to identify and disrupt attack paths before they escalate into sustained access or operational impact.

# Systemic Weaknesses Enabled Outsized, Ecosystem-Wide Impact

While most cloud intrusions in 2025 impacted organizations individually, the year was also characterized by a smaller number of breaches with outsized, ecosystem-wide impact. In these cases, threat actors focused less on targeting single environments and more on identifying shared technologies, dependencies, and platforms used across many organizations.

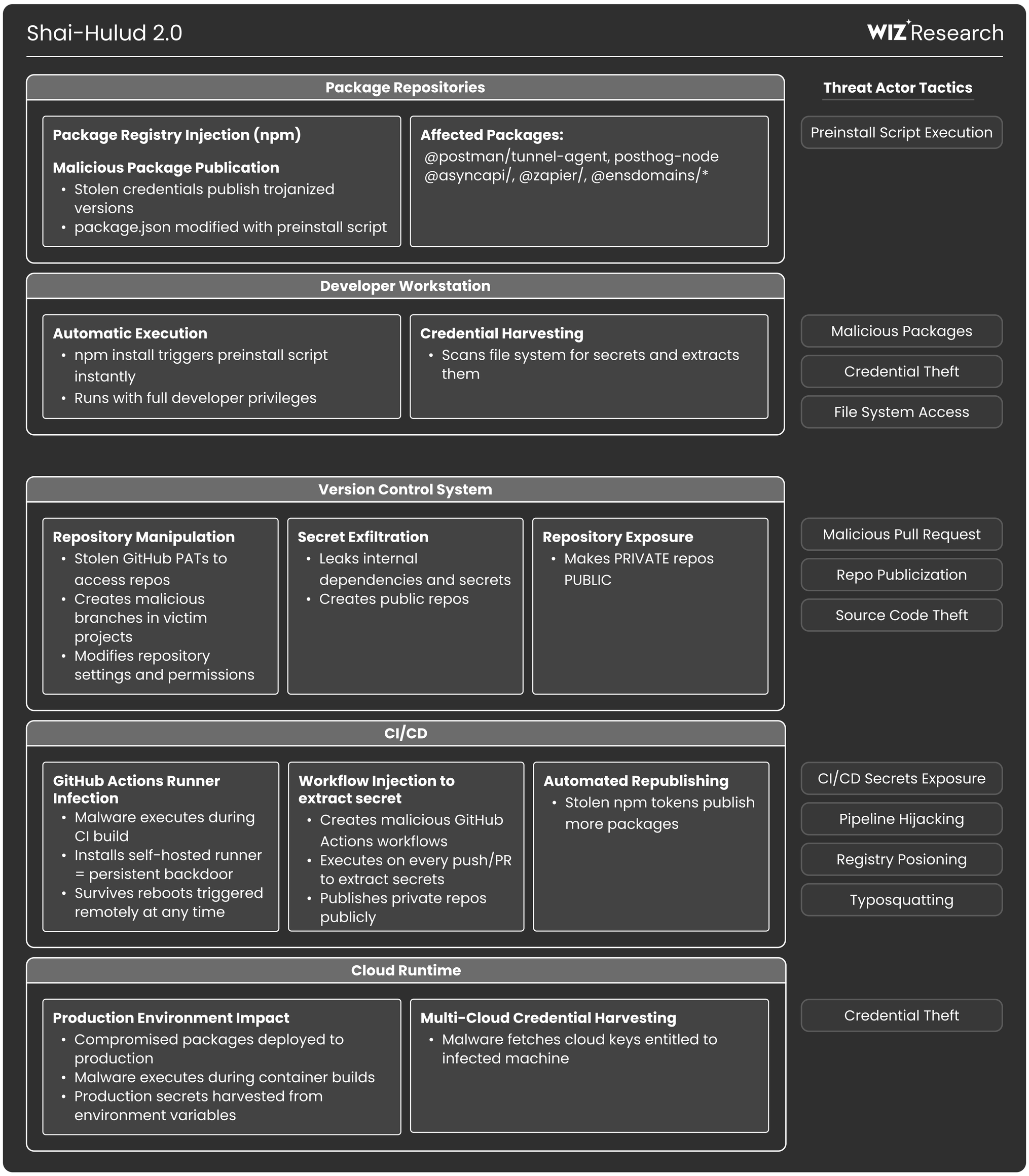



When widely adopted, internet-facing software or integrations contain a weakness, that weakness can affect many environments at once. Threat actors abused trusted vectors and credentials stored at outside vendors to take advantage of systemic points of shared trust.

## For Example

### 1 Shai Hulud: Supply Chain Access at Scale

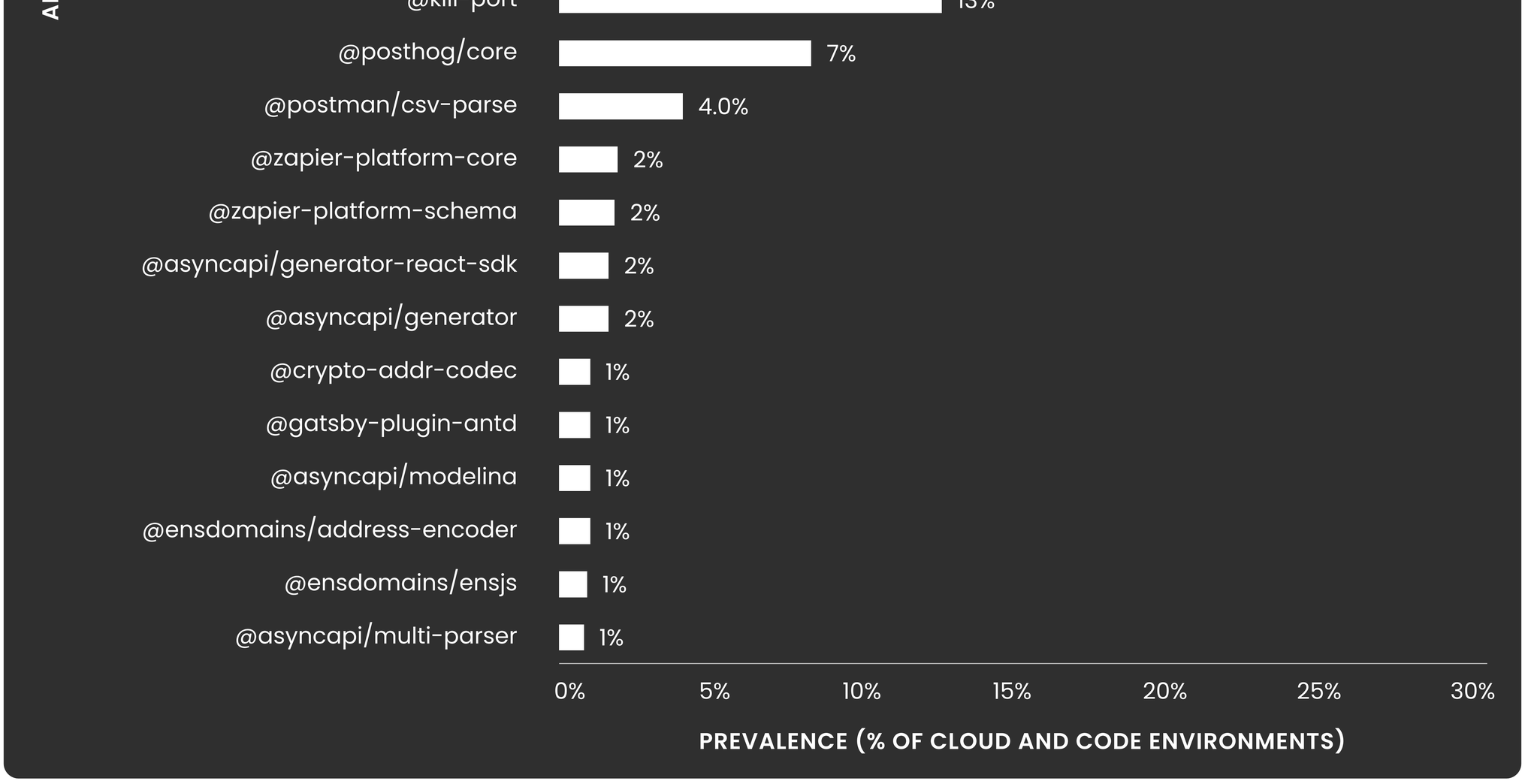
The Shai Hulud supply chain attacks were some of the most consequential cloud-related incidents of the year due to their downstream impact across a large number of organizations. Wiz Research analyzed multiple phases of the campaign, including its initial discovery, its subsequent evolution, and aftermath.



Rather than exploiting individual cloud environments directly, attackers compromised a widely used npm package and leveraged inherited trust within the software supply chain. Malicious behavior propagated downstream through legitimate development workflows, impacting organizations that had not been directly targeted.

While the underlying access strategy of abusing trusted dependencies was well established, Wiz Research found that the scale of impact was driven by how broadly that dependency was embedded across the ecosystem. A single compromise translated into exposure across many environments, turning a localized intrusion into a multi-organization security event.

The following graph shows the percent of Wiz-monitoried environments where the listed packages were found.



## 2 s1ngularity: Abuse of Shared Tooling and Automation

The s1ngularity attack followed a similar pattern, targeting shared tooling rather than individual cloud environments. Wiz Research investigated the incident and its aftermath, documenting how attackers abused trusted automation components to propagate malicious behavior.

In this case, attackers relied on legitimate build and automation workflows to distribute access. Organizations affected by s1ngularity were exposed not because of unique misconfigurations, but because they relied on the same tooling and integration patterns.

As with Shai Hulud, attackers relied on familiar techniques that operated through trusted systems and credentials. This enabled them to reduce repeated exploitation, and scale impact across many targets.

## 3 Internet facing zero-days enable massive compromise

Multiple times this year, vulnerabilities in internet facing technologies forced security teams around the world to scramble, as threat actors quickly weaponized exploits and began gaining access to networks. The biggest of these was the React2Shell (CVE-2025-55182) vulnerability. This vulnerability in React allowed attackers to send a specially crafted message and execute arbitrary code.

On December 3, the first few hours following the release of the vulnerability were relatively quiet, with some state backed groups identified testing possible exploits. However, the next day when a public proof of concept was released, Wiz identified dozens of exploitation campaigns within hours. Over the next week, over 60 distinct campaigns would exploit this vulnerability for a variety of schemes.

The most common was crypto mining, representing 40% of the campaigns. This was followed by campaigns that installed backdoors at 21% and campaigns that simply exfiltrated credentials comprising another 18%. These campaigns varied widely in sophistication, ranging from complex, state-backed operations to basic AI-generated scripts. However, the speed and breadth meant that anyone with a vulnerable system was likely affected.

## 4 Trusted Relationship Abuse Beyond Large-Scale Campaigns

Wiz Research has observed the same techniques used in more targeted intrusions as well. In investigations into GitHub personal access token (PAT) abuse, threat actors leveraged stolen or overly permissive tokens to operate through GitHub's control plane, inheriting access to repositories, CI workflows, and connected cloud environments without exploiting vulnerabilities directly.

These incidents demonstrate that abuse of trusted relationships is not limited to large-scale, opportunistic supply chain attacks. The same approach can be applied selectively against individual organizations by targeting the identities, credentials, and integrations that underpin modern development and deployment workflows.

### 5 Abuse of OAuth Tokens in Widely Integrated SaaS Platforms

In several 2025 incidents analyzed by Wiz Research, attackers leveraged stolen OAuth tokens and API credentials to access widely used SaaS platforms and connected cloud services. By abusing valid authentication material rather than exploiting platform vulnerabilities, attackers operated through legitimate APIs and inherited access to downstream systems.

Because many organizations rely on similar integration models, the compromise of a single entity created shared exposure across multiple environments, extending impact well beyond the initially affected organization.

These systemic breaches highlight a different risk dynamic than the individual intrusions described earlier in this report. They demonstrate how shared infrastructure, dependencies, and platforms can magnify the consequences of otherwise well-understood attack strategies, making visibility into supply chains, integrations, and inherited trust increasingly critical to cloud security.

## AI Expanded the Attack Surface and Accelerated Familiar Techniques

AI did not necessarily introduce fundamentally new cloud attack techniques in 2025. Instead, it shaped how familiar cloud intrusions unfolded by expanding the exposure of AI-driven infrastructure and accelerating attacker workflows through AI-assisted tooling.

### 1 AI Adoption Expanded the Attack Surface

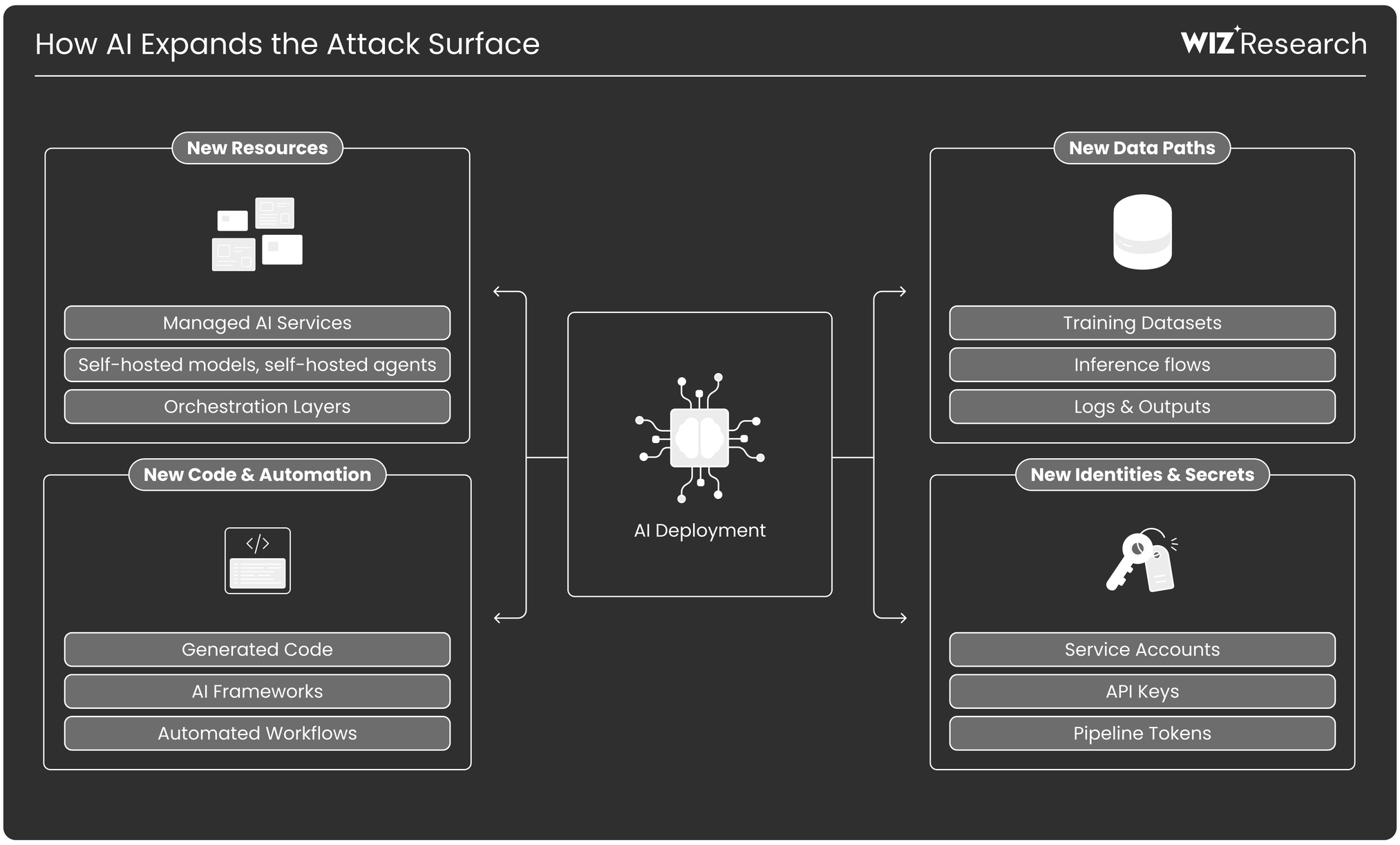
AI adoption is now widespread across cloud environments. Wiz Research from early 2025 found that **over 85% of organizations are using some form of AI**, whether through managed services or self-hosted deployments. Survey data collected in late 2024 for the AI Security Readiness report similarly shows that **87% of respondents were already using AI services.**

This growth matters because many organizations report they are adopting AI faster than they are adapting security practices:

- **25% of respondents said they do not know which AI services are running in their environment,** indicating gaps in visibility and governance.
- **31% cited a lack of AI security expertise** as their top challenge.
- While traditional controls such as endpoint and vulnerability management remain common, **only 13% reported using AI-specific posture management (AI-SPM).**

As AI systems proliferated, they introduced new services, pipelines, identities, and integrations that were often tightly connected to sensitive data, privileged identities, and high-value compute resources.

Wiz Research documented multiple cases where familiar cloud risks such as misconfigurations, exposed credentials, and excessive permissions were observed within AI-related systems.

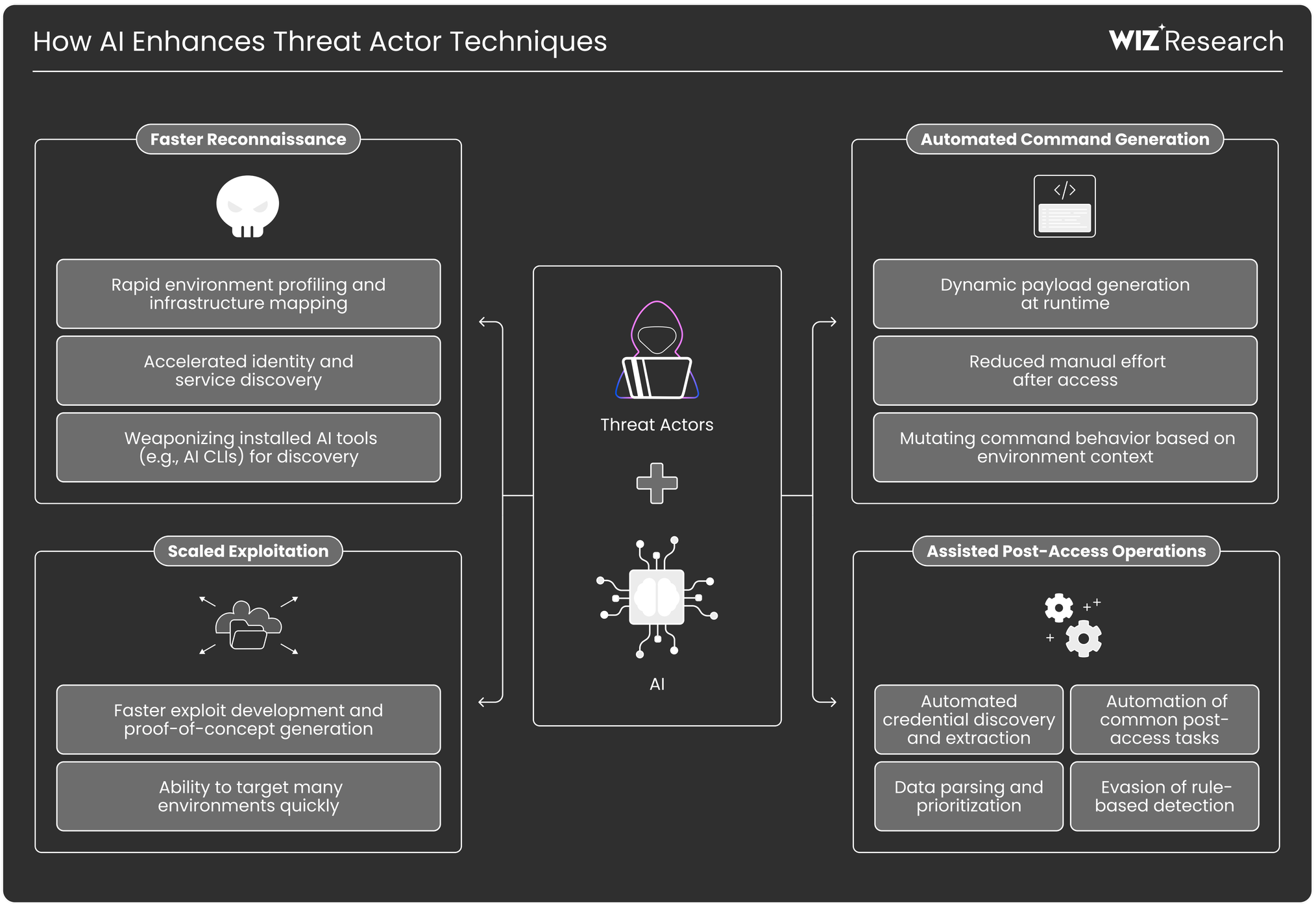



**For example:**

- **Widespread leakage of AI-related secrets in development ecosystems:** Wiz's investigation into the Forbes AI 50 found that 65% of leading private AI companies had leaked verified secrets such as API keys, tokens, and credentials on GitHub. Many of these secrets were hidden in commits, deleted forks, and developer repositories, creating a large credential surface that attackers can leverage for automation or tooling misuse.
- **AI-accelerated development is producing real, repeatable app security failures:** Wiz Research identified critical flaws in a vibe-coding platform enabling unauthorized access to private apps, and found systemic security issues affecting a meaningful portion of vibe-coded applications.
- **AI pipelines introduce shared-trust risk through credentials, automation, and shared infrastructure:** Wiz Research has shown how AI build and inference workflows rely on service accounts, API tokens, and shared automation components. When these credentials are compromised, access can propagate across environments, extending the impact of credential theft or misconfiguration beyond a single system.
- **Exposed AI orchestration surfaces are being actively exploited:** Wiz documented a cryptomining campaign targeting Ray clusters. The actors used Ray's native scheduling and orchestration features to propagate malicious workloads across GPU-backed nodes.

Across these cases, AI did not create new classes of vulnerability. Instead, it expanded the number of places where well-known risks could appear, often closer to sensitive data and compute-intensive resources.

## 2 AI Accelerated Familiar Attack Techniques




Over the past year, threat actors also incorporated AI tooling into their operations in a variety of ways.

**For example:**

- **AI-assisted malware execution.** Wiz Research analyzed malware such as LameHug, a strain observed in July 2025 that leverages prompts to large language models to generate system commands on compromised machines. Instead of relying on static hardcoded instructions, LameHug sends encoded prompts to an external AI model to dynamically produce reconnaissance and data-gathering commands at runtime.
- **Abuse of AI command-line tools in real incidents**. In the s1ngularity supply chain attack, malicious packages introduced into the popular Nx build system leveraged installed AI CLI tools such as Claude, Gemini, and Q to aid in reconnaissance and credential harvesting as part of post-compromise operations. This approach illustrates how familiar developer utilities can be repurposed by attackers once a foothold exists.
- **AI-assisted reconnaissance and automation after access.** Wiz's s1ngularity's Aftermath analysis details how attackers used AI-linked tooling to automate the discovery of credentials and environment data during both the initial and follow-on phases of the compromise, reducing manual effort and accelerating traditional exfiltration techniques.

In addition to real-world incidents, Wiz's zeroday.cloud event provided an illustrative example of how AI-assisted tooling can lower the effort required to discover and operationalize vulnerabilities. During the event, security researchers used AI to identify and reason about flaws in widely used technologies over short timeframes. In one case, the winning researcher identified multiple zero-day vulnerabilities using AI without direct human involvement, demonstrating how tasks that have traditionally required significant time and specialized expertise can be compressed dramatically in a controlled research setting.

While zeroday.cloud reflects a research competition rather than threat actor activity in the wild, it highlights how AI can support vulnerability discovery and exploitation workflows without introducing new attack classes. Similar dynamics were observed in real-world incidents, such as MongoBleed (CVE-2025-14847), where an AI-generated proof of concept was produced within minutes and later used in active exploitation. This demonstrated how AI can accelerate the operationalization of known vulnerability patterns once flaws are disclosed.

These observations suggest that AI's primary impact on cloud security in 2025 was **contextual rather than transformational.** Understanding where AI systems intersect with identities, data paths, and cloud services remains critical, but the foundational risks defenders must manage have not fundamentally changed.

# What Defenders Can Do

Effective defense against the threats we've discussed does not require entirely new strategies, but it does require applying core controls with greater speed, context, and visibility.

Based on how threat actor activity unfolded across the reporting period, defenders should prioritize the following actions:

- ☑ **Reduce externally reachable exposure before threat actors arrive.**
  Vulnerabilities, exposed secrets, and misconfigurations accounted for the majority of initial access we observed in 2025. Defenders should focus on identifying which assets are externally reachable and which risks are exploitable from the outside. From there, they should prioritize remediation where those issues are connected to sensitive identities, data, or workloads. Continuous visibility into exposure and attack paths can help teams focus remediation on risks that are realistically exploitable.

- ☑ **Treat pre-compromise reconnaissance as a detection opportunity.**
  Threat actors consistently performed scanning, discovery, and access preparation before establishing durable access. Visibility into reconnaissance activity, dependency mapping, and anomalous discovery behavior can provide early warning and enable disruption before persistence and impact occur.

- ☑ **Harden identities and monitor control-plane activity after access.**
  Once access was established, threat actors rapidly pivoted to cloud identities and services rather than relying on traditional lateral movement. Monitoring identity changes, privilege expansion, and API activity across cloud control planes and SaaS integrations is critical to detecting post-compromise abuse of legitimate access.

- **Limit blast radius from shared trust and supply chain dependencies.**
  Incidents involving compromised packages, CI systems, SaaS integrations, and automation workflows demonstrated how inherited trust can extend impact beyond a single environment. Defenders should maintain visibility into trusted relationships across development pipelines, third-party services, and identity federations, and correlate these relationships with exposure and identity risk to reduce downstream impact.
- **Apply production-grade security controls to AI systems and workflows.**
  AI-driven infrastructure introduces new services, identities, and data paths that often intersect directly with high-value resources. Defenders should inventory AI workloads and pipelines, apply consistent access controls and configuration standards, and continuously monitor how these systems connect to cloud identities and data.

# Conclusion

In our observations, threat actor activity targeting cloud environments during 2025 was defined less by new techniques than by increased scale and interconnectedness. Threat actors continued to rely on familiar weaknesses, but applied them across shared infrastructure, automation, and AI-assisted workflows. When access was observed or threat actors attempted to expand control, activity consistently focused on cloud identities, control planes, and high-value services, reflecting typical escalation behavior in cloud environments.

For defenders, strong fundamentals remain essential, but they must be applied with greater context. Organizations that maintain visibility into exposure, identity relationships, and how risk propagates across cloud, development, and AI environments are better positioned to detect and disrupt intrusion activity before it escalates into meaningful impact. Attack surface management plays a critical role in this effort by helping teams identify and prioritize externally reachable risks.

# Methodology

This report analyzes cloud intrusion patterns and threat actor behavior observed between February and December 2025 and compares them with patterns identified in the prior reporting period (September–December 2024). In this report, we use the term “threat actor” to describe any entity, whether human or automated, that attempts to compromise cloud environments.

### The analysis draws on three primary sources:

1. **Publicly documented cloud security incidents,** including supply chain attacks, vulnerability exploitation, and ecosystem-wide events, used to characterize how cloud attacks and breaches occur across the industry. These incidents are analyzed as part of Wiz’s Cloud Threat Landscape research.
2. **Aggregated Wiz detection telemetry mapped to the MITRE ATT&CK framework,** which provides visibility into attempted and detected malicious activity during access-preparation and post-access phases in cloud environments. Post-access activity refers to threat actor actions observed after an initial foothold and does not imply persistence or operational impact.
3. **Investigations conducted by** Wiz Research and Customer Incident Response teams, which provide insight into how cloud intrusions unfold across a curated set of real-world cases.

Together, these sources are used to identify recurring threat actor behaviors across the intrusion lifecycle, including initial access methods, access-preparation activity, and post-access techniques. Detection-based activity is analyzed based on the prevalence of malicious action types per tenant over time, rather than raw alert volume, to highlight common threat actor behaviors without implying successful compromise.

Detection severity levels (such as high and critical) are assigned based on the potential risk associated with the observed behavior, including interaction with cloud control planes, identities, or sensitive resources. High- and critical-severity detections reflect actions that could enable escalation or impact if left unaddressed and do not always mean successful compromise, persistence, or operational disruption.

The environments informing this analysis span a broad cross-section of large enterprise cloud deployments across industries and cloud providers. While the findings do not represent the full scope of global cloud activity, the consistency of patterns observed across data sources and reporting periods supports their relevance to modern cloud threat dynamics.