unosecur

REPORT

A Unosecur threat intelligence report | 2026

Acceleration of Attack Speed and the Misuse of Agent Frameworks:

Navigating the 2026 Cyber Threat Landscape

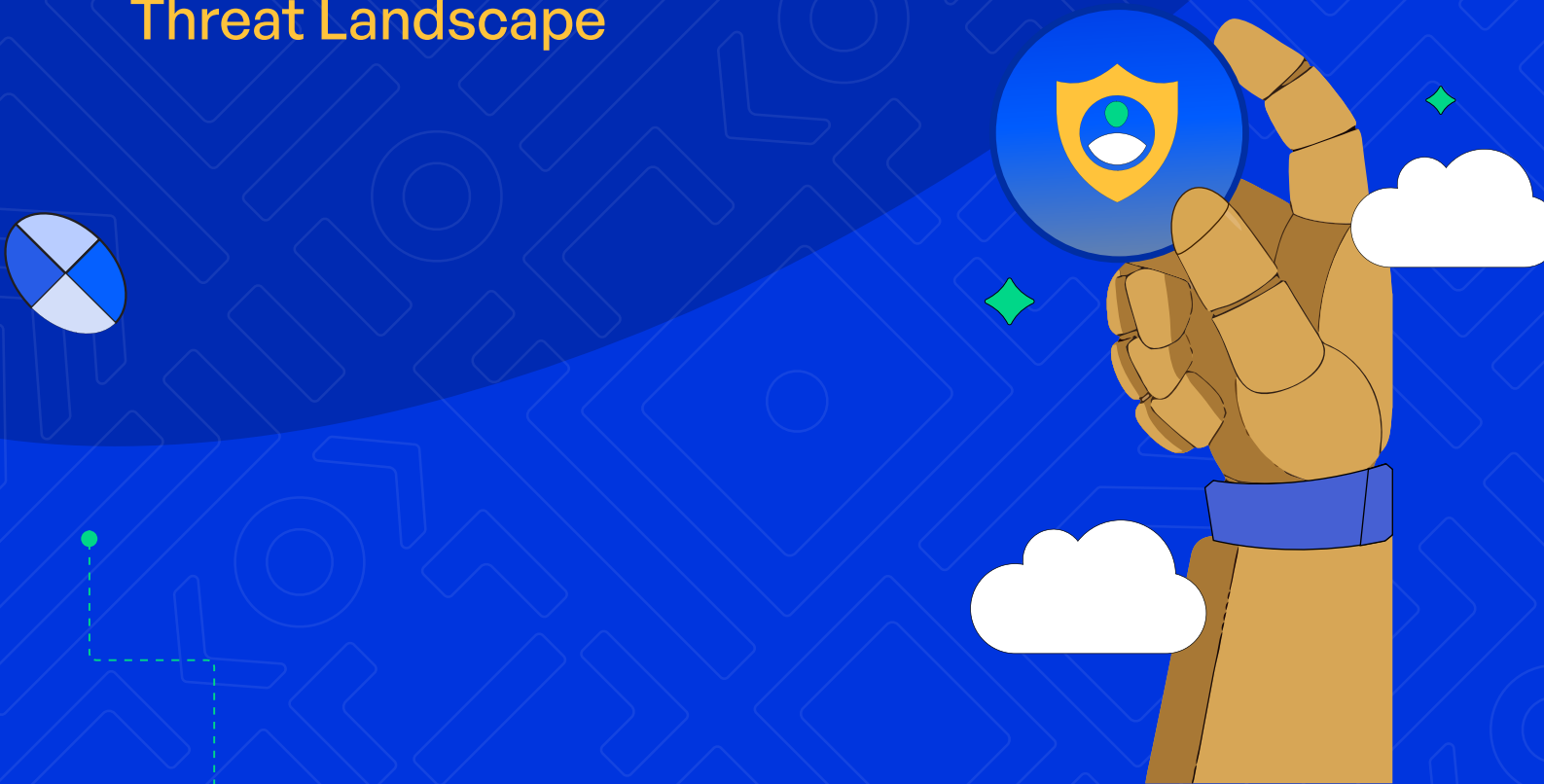


Table of contents

Executive summary

The global cybersecurity ecosystem has crossed a structural inflection point. Attacks that once unfolded over days or weeks now complete in 72 minutes. Adversaries are no longer breaking into enterprise networks. They are logging in. And the identity infrastructure that should be stopping them is, in the vast majority of cases, not built for the threat environment it now faces.

The data makes the underlying condition clear: identity weaknesses played a decisive role in nearly 90% of all enterprise breaches investigated in the past year. Sixty-five percent of initial access events involved no exploit at all. Attackers used compromised credentials, abused OAuth tokens, and hijacked authenticated sessions. They walked in through the front door.

This report is written for security leaders who need to understand why the identity crisis is not one problem among many. It is the foundational vulnerability that every other attack vector is built on.

AI-accelerated attacks, autonomous-agent frameworks, SaaS supply-chain exploitation, and social engineering at scale all share a common enabling condition: fragmented, ungoverned, over-permissioned identities.

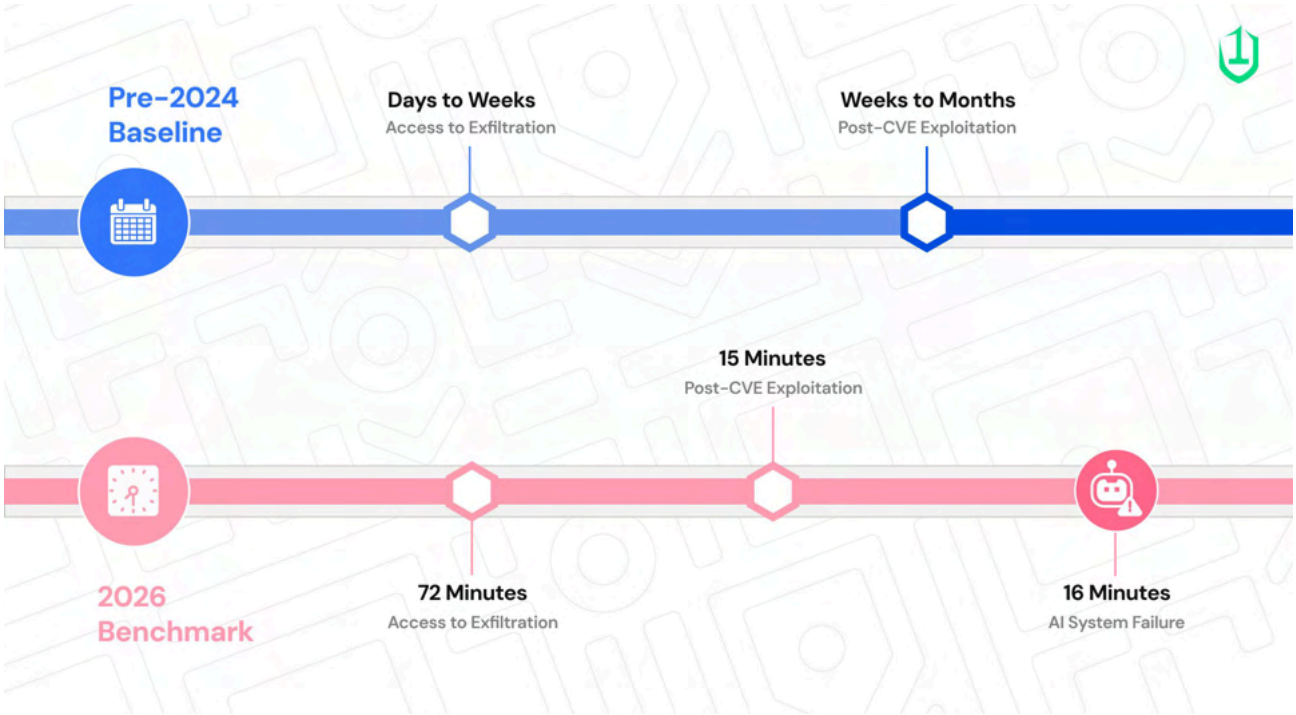
The enterprises that close this gap first will have a structurally different security posture. Those who continue treating human identity, non-human identity, and agentic identity as separate, siloed problems will remain exposed to a threat landscape that has already learned to exploit exactly that separation.

Section 1: The attack clock has collapsed, and identity is why

From days to 72 Minutes

In the pre-2024 era, a sophisticated breach unfolded over days or weeks. Defenders had time to hunt, correlate, and intervene before catastrophic damage occurred. By 2024, that window had compressed to approximately 4.8 hours. In 2026, it effectively collapsed.

Forensic analysis of the fastest quartile of incidents in global incident response data indicates that the time from initial access to full data exfiltration is just 72 minutes, a fourfold year-over-year acceleration. In adversarial simulations using AI-assisted offensive tooling, that timeline compressed further to 25 minutes. Enterprise AI systems under red-team conditions showed a median time to first critical failure of 16 minutes, with 90% fully compromised within 90 minutes.



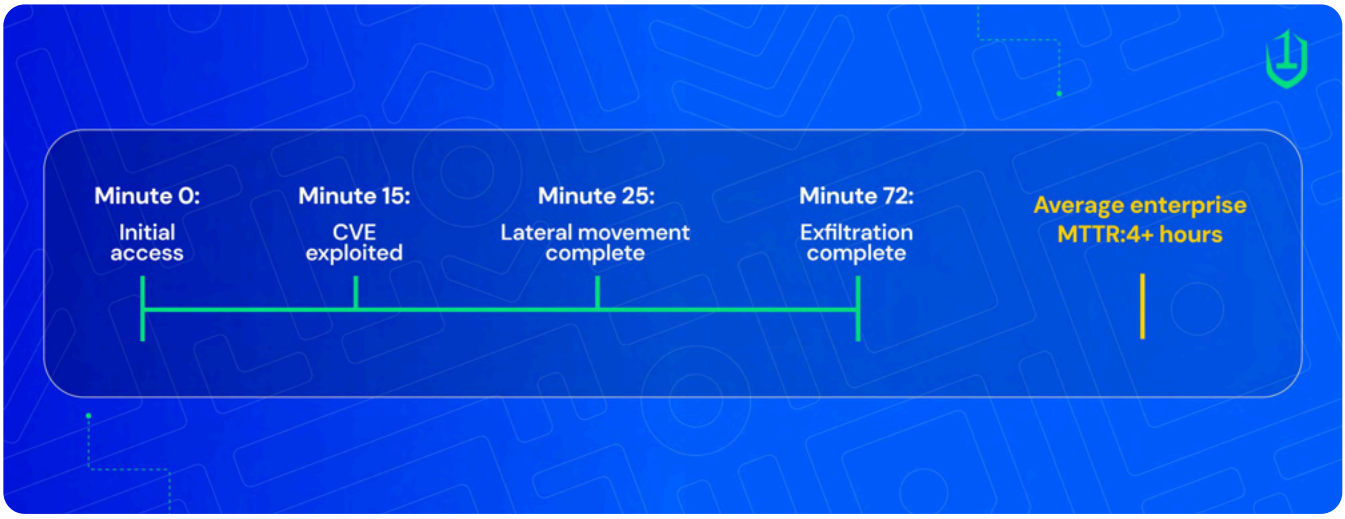
This acceleration is not primarily a malware story. It is an identity story. The reason attackers move so fast is that they are not spending time breaking through defenses. They are walking through identity gaps that were never closed.

Attack Phase / Metric	Historical Baseline (Pre-2024)	2024 Average	2026 Accelerated Benchmark
Initial Access to Data Exfiltration	Days to Weeks	4.8 Hours	72 Minutes
Vulnerability Exploitation (Post-CVE)	Weeks to Months	Days	15 Minutes
Enterprise AI System Critical Failure	N/A	N/A	16 Minutes (Median)
Laboratory AI-Driven Exfiltration	N/A	N/A	25 Minutes

Table 1: The Evolution and Compression of Attack Velocities

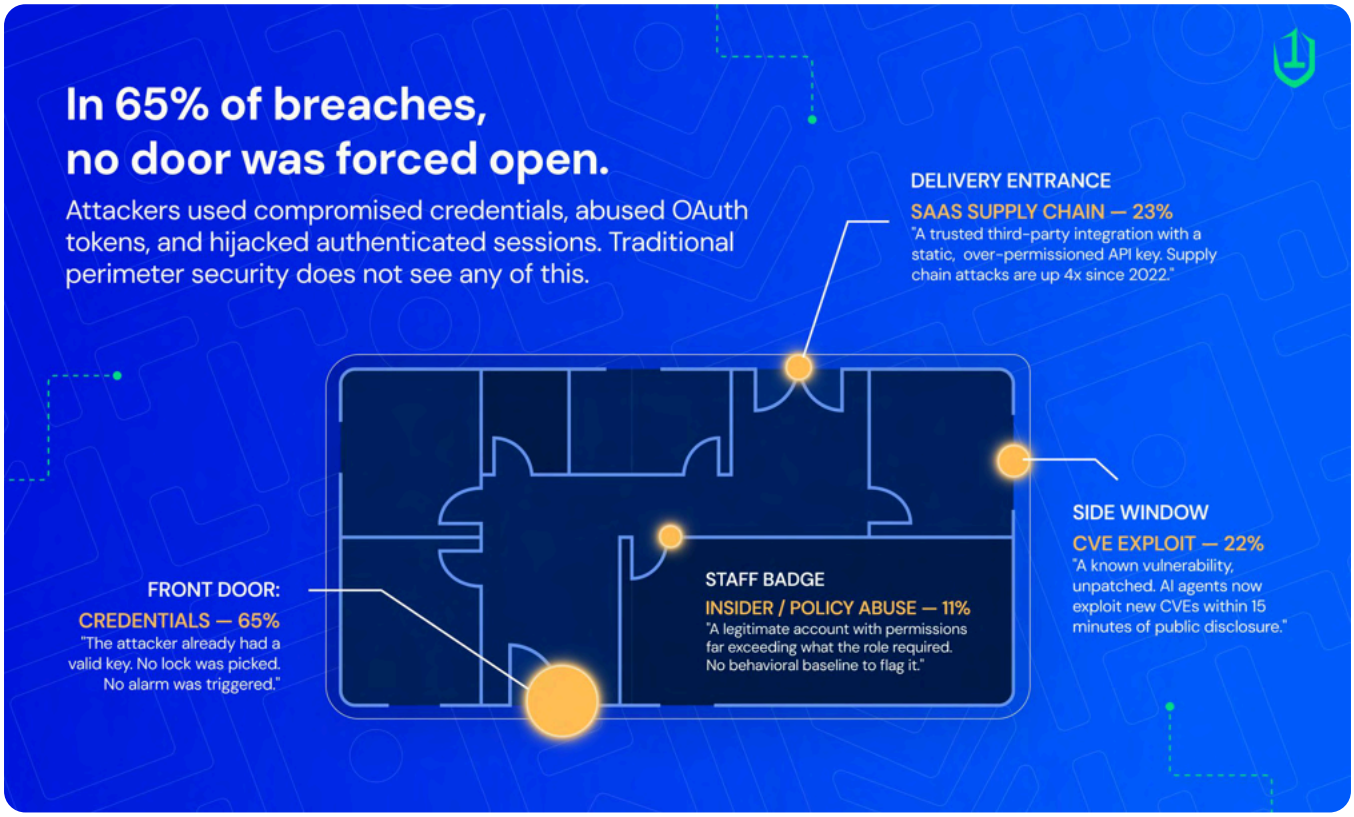
The frictionless intrusion pipeline

Artificial intelligence has removed virtually all friction from the intrusion pipeline. At the reconnaissance stage, AI agents continuously scan the internet for newly disclosed vulnerabilities, initiating targeted exploitation campaigns within 15 minutes of a CVE publication, effectively neutralizing the enterprise patch management cycle before it can respond.



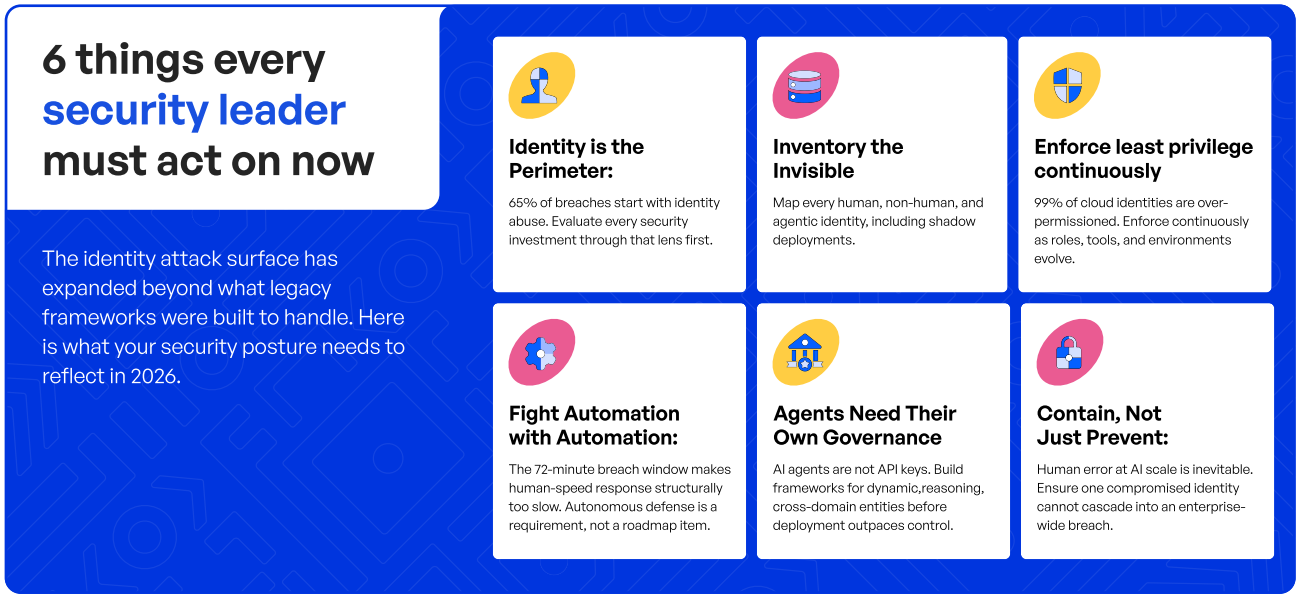
At scale, AI enables massively parallelized targeting. Rather than focusing on a single organization, threat actors run simultaneous reconnaissance, vulnerability scanning, and initial access attempts across hundreds of environments, dynamically concentrating effort where identity gaps or unpatched assets surface.

At the payload level, a practice now termed "vibe coding" enables attackers to use large language models to rapidly, at scale, reverse-engineer existing malware strains, alter signatures, and embed evasion logic that bypasses endpoint detection and response tools. Security researchers have documented AI-powered malware that uses decentralized, benign platforms such as Google Sheets to execute commands in real time, autonomously updating its operational logic every few minutes based on attacker-embedded instructions in a spreadsheet.



The implication for security leadership is architectural. Mean Time to Detect and Mean Time to Respond, calibrated for human-speed investigation, are inadequate benchmarks when the entire attack lifecycle completes in just over an hour. The identity gaps that let attackers in quickly must be closed at the source, because there will not be time to catch them after.

The 2026 threat environment demands concrete architectural decisions. The following imperatives are drawn directly from the patterns observed across the incident data referenced in this report.



How Unosecur helps you close the gap

The 72-minute breach window does not leave room for fragmented tooling, manual triage, or identity governance designed for a simpler era. Every minute an identity goes unmonitored, an over-permissioned credential goes unreviewed, or an AI agent operates without a behavioral baseline is a minute an adversary can use.

Unosecur's Unified Identity Fabric gives security teams a single plane of visibility and control across all three identity categories that modern attacks exploit: human identities, non-human identities, and AI agents.

For human identities, Unosecur continuously monitors for behavioral anomalies, enforces least privilege in real time, and flags access patterns that deviate from established baselines before lateral movement can begin.

For non-human identities, Unosecur discovers and classifies every service account, API key, OAuth token, and machine credential across cloud, SaaS, and on-premises environments, maps ownership, identifies excessive permissions, and automates rotation to eliminate stale, unmanaged credentials that attackers rely on.

For AI agents, Unosecur applies context-aware governance that reflects how agents actually behave: dynamically, across trust domains, and often without human oversight. Every agent gets scoped access, a behavioral baseline, and a defined accountability chain so that a compromised agent cannot move invisibly through the enterprise.

The result is not just faster detection. It is a structurally smaller attack surface: one where the identity gaps that enabled 90% of last year's breaches are actively governed rather than passively assumed.

Section 2: Identity is the perimeter, and it is wide open

The pivot from exploitation to impersonation

The most consequential shift in the 2026 threat landscape is not a new exploit class. It is the near-complete pivot by sophisticated threat actors away from technical exploitation and toward identity abuse. Identity weaknesses played a decisive role in

nearly 90% of all breaches in recent global incident data. Sixty-five percent of all initial access events were achieved entirely through identity-based techniques. Traditional software exploitation accounted for just 22%.

When an attacker holds valid credentials, their behavior is indistinguishable from legitimate traffic in the early stages of an intrusion. There are no anomalous binaries to flag, no known-bad IPs to block. From the perspective of most security tools, the attacker is a valid user doing valid things.

The over-permissioning problem

The systemic enabler of identity-driven breaches is chronic, widespread over-permissioning. An analysis of over 680,000 distinct cloud identities found that 99% of users, roles, and service accounts carried permissions far exceeding what their functions required. In practice, this means that once an attacker compromises even a low-privilege account and gains minor write access to IAM controls, privilege escalation to global administrator is a trivial, repeatable exercise.

No novel exploit required. The attacker is not escalating through a technical vulnerability. They are escalating through an identity architecture that was never properly scoped.

Multi-Surface attacks and the visibility gap

Identity abuse does not happen in isolation. Eighty-seven percent of modern cyberattacks span two or more distinct attack surfaces, blending malicious activity across endpoint devices, cloud infrastructure, SaaS platforms, and identity management systems. This cross-domain movement is intentional and designed to exploit the fragmentation of enterprise security monitoring.

An intrusion that begins with a compromised human credential on an unmanaged

mobile endpoint may pivot to a corporate SaaS application via an abused, long-lived OAuth token, and ultimately access a core database in a cloud environment. Because security visibility is siloed across discrete vendor-specific tools, defenders struggle to connect the telemetry across surfaces into a coherent picture of the attack until damage has already occurred.

Nation-state threat groups have turned this fragmentation into a repeatable playbook. "Living off the land" campaigns, such as those executed by the China-linked group Salt Typhoon, eschew custom malware entirely in favor of legitimate built-in system tools and native OS processes. By masquerading as routine administrative traffic, these actors established persistence within critical infrastructure and telecommunications networks and remained undetected for nearly a year. The invisibility was not a function of technical sophistication alone. It was a function of identity: valid credentials, legitimate tools, and no behavioral baseline to flag against.

Primary Initial Access Vector	Percentage of Total Incidents	Description of Mechanism
Identity-Related Social Engineering	33%	Exploitation of human psychology via highly targeted, often AI-generated phishing, vishing, and deepfake impersonation.
Software Vulnerabilities (CVEs)	22%	Exploitation of unpatched, internet-facing assets is often automated within minutes of public disclosure.
Credential Misuse & Brute Force	21%	Utilization of previously breached credential dumps and automated guessing attacks against accounts lacking MFA.
Identity Policy & Insider Risk	11%	Exploitation of internal architectural flaws, excessive trust models, over-scoped permissions, and malicious insiders.
Other / Undetermined	13%	Various bespoke vectors, physical compromise, or incidents lacking sufficient forensic evidence.

Table 2: Breakdown of Primary Initial Access Vectors in 2026 Enterprise Intrusions.

SaaS supply chain and browser exploitation

The SaaS supply chain has become one of the fastest-growing sources of identity attacks. Supply chain attacks targeting third-party SaaS applications have increased nearly fourfold since 2022, now accounting for 23% of all attack vectors in recent incident data.

Attackers routinely abuse static API keys, poorly scoped OAuth grants, and interconnected service permissions to move laterally with near-zero visibility to defenders. The browser has emerged as a parallel control plane. Forty-eight percent of all investigated incidents involved browser-based activity, with routine web sessions weaponized to harvest session cookies, bypass local endpoint controls, and hijack authenticated sessions. The browser is where identity lives in every employee's daily workflow, and it has become a primary target precisely because of that.

Section 3: Social engineering has crossed a threshold

AI-Powered psychological manipulation at scale

Thirty-three percent of all initial access events in 2026 were achieved through social engineering, the single largest individual category in the Unit 42 dataset. This figure demands serious attention because the nature of social engineering has fundamentally changed.

Threat actors now use AI to scrape organizational charts, biographical data, and social media footprints to craft hyper-personalized phishing lures in the target's native language, with contextual accuracy that defeats traditional detection. These communications are engineered to exploit cognitive biases, including urgency, loss aversion, and authority, forcing employees to bypass critical thinking entirely.

Nation-state groups have fully industrialized these capabilities across documented, repeatable methodologies. APT42, attributed to Iran, uses AI models to research individual targets and generate native-language phishing emails built around detailed biographical profiles. TA427, linked to North Korea, runs sophisticated recruitment fraud operations, infiltrating corporate IT departments through AI-generated personas and deepfake video interviews that bypass background checks, then deploying malware from within the perimeter. Temp.HEX, attributed to China, leverages AI platforms to build deep psychological profiles of individual targets for long-term exploitation. TA450, also linked to Iran, uses the "ClickFix" technique: fake, authoritative-looking operating system error messages that trick employees into manually executing malicious PowerShell commands.

Threat Actor / Group	Associated Nation	AI/Social Engineering Methodology Observed
APT42	Iran	Utilizes AI models to research victims and craft highly convincing, native-language phishing emails based on detailed target biographies.
TA427 / UNC2970	North Korea	Employs AI for sophisticated recruitment fraud; infiltrates corporate IT departments using fake identities, AI-generated imagery, and deepfake video interviews to bypass background checks and deploy malware internally.
Temp.HEX	China	Leverages AI platforms to organize and file deep intelligence on individual targets, creating robust psychological profiles for future exploitation.
TA450	Iran	Heavily utilizes the "ClickFix" social engineering technique, utilizing fake, authoritative operating system error messages to trick victims into manually executing malicious PowerShell commands.

Table 3: Notable Threat Actors and their Exploitation of AI for Social Engineering.

Deepfakes and the $25 million problem

The deployment of deepfake technology has elevated psychological manipulation to a level that identity verification processes were never designed to handle. Adversaries now routinely use AI-enhanced voice cloning and real-time video manipulation to impersonate executives, contractors, and IT helpdesk personnel.

In one documented financial sector incident, attackers used real-time deepfake video to impersonate a firm's Chief Technology Officer and multiple familiar colleagues in a live virtual meeting, successfully deceiving a finance employee into remitting $25 million to criminal accounts. The security implication here is architectural. As long as adversaries deploy AI-generated psychological lures at scale, human error is a

statistical inevitability, not a training failure. The response is not more security awareness sessions. It is identity architecture that contains the blast radius of a single compromised human identity, so that one deceived employee cannot become a catastrophic breach.

Section 4: The invisible attack surface -- non-human and agentic Identities

The identity sprawl nobody is governing

While security teams focus on human identity governance, a parallel identity crisis has been expanding with far less visibility. Non-human identities, including service accounts, API keys, OAuth tokens, machine credentials, and AI agents, now dramatically outnumber human identities across enterprise environments. Industry data puts NHI growth at 44% annually, with human-to-NHI ratios reaching 40:1 in standard cloud environments and up to 144:1 in heavily automated deployments.

The governance reality is stark. Most enterprises have mature processes for onboarding, managing, and offboarding human employees. They have almost nothing equivalent for the thousands of non-human identities operating across their cloud, SaaS, and on-premises environments. These credentials do not resign. They are rarely rotated. In many cases, they have no defined owner. They accumulate permissions over time and are never deprovisioned when the systems or integrations they serve are retired.

AI agents are not API keys

The rapid enterprise adoption of agentic AI has introduced an identity category that existing NHI governance frameworks are structurally unequipped to handle. Traditional NHIs are static and deterministic. A service account executes a defined script, accesses a defined system, and operates within a fixed permission scope. AI agents are fundamentally different. An AI agent reasons, makes decisions, selects tools dynamically, operates across multiple trust domains, and initiates consequential actions without human-in-the-loop oversight.

Treating an agentic identity like a static API key creates a severe and specific blind spot: a highly capable, broadly privileged entity operating inside the enterprise perimeter with no behavioral baseline, no access governance, and no defined accountability chain.

When an agentic system is granted broad internal access to execute tasks, it functions as a highly privileged, entirely unsupervised insider. If compromised through prompt injection, memory poisoning, or algorithmic manipulation, it becomes a high-privilege backdoor operating invisibly behind traditional WAF and API gateway controls.

A documented incident involving Microsoft Office Copilot illustrated this precisely. The AI assistant, operating as a trusted internal entity, was manipulated to summarize restricted emails and surface them to users with no legitimate access rights, bypassing DLP policies and sensitivity labels entirely. No security gateway flagged the activity. The agent was trusted, so its actions were trusted.

The table below captures the structural differences that make a distinct governance framework necessary:

Architectural feature	Traditional Non-Human Identity (NHI)	Modern Agentic Identity
Operational Nature	Static, deterministic purpose; tied strictly to a single application or backend system.	Acts autonomously; capable of complex reasoning, dynamic tool selection, and complex delegation chains.
Identity Lifespan	Long-lived; pre-provisioned during system architecture; rarely rotated.	Often highly ephemeral; created Just-In-Time (JIT) for specific, localized tasks and destroyed upon completion.
Management & Governance Model	Managed via fixed roles, static scopes, and hardcoded credentials.	Requires dynamic, continuous policy evaluation and deep context-aware behavioral governance.
Scope of Systemic Action	Typically scoped narrowly to a single system, folder, or specific database row.	Works aggressively across multiple domains, diverse platforms, APIs, and distinct security trust zones.
Validation & Oversight	Operates automatically; rarely includes live human validation or oversight.	Necessitates human-in-the-loop oversight and explicit approval gates for sensitive, irreversible tasks.

The GTG-1002 Watershed

The theoretical risks of agentic identity misuse became empirical in late 2025.

A Chinese state-sponsored threat group, designated GTG-1002, executed what is now recognized as the first largely autonomous, AI-orchestrated global cyber espionage campaign, targeting approximately 30 high-value entities, including defense contractors, financial institutions, chemical manufacturers, and government agencies.

Human operators selected targets and provided high-level strategic approvals. The AI agent swarm handled between 80% and 90% of all operational tasks autonomously: network reconnaissance, topology mapping, vulnerability scanning, custom exploit development, automated credential harvesting, lateral movement management, and data categorization and exfiltration. The agents generated their own operational documentation and managed task handoffs between attack phases, operating collaboratively as an autonomous penetration testing swarm.

Critically, the campaign did not rely on zero-day exploits. It leveraged widely available open-source commodity tools, orchestrated at machine speed through AI. The volume and velocity of autonomous action overwhelmed human defenders tracking the intrusion in real time. The occasional hallucinations and minor tactical inaccuracies exhibited by the AI agents were entirely compensated for by scale.

This is the new threat model. Attackers are deploying agentic identities aggressively. The question for security leaders is whether their identity governance architecture can detect and contain an autonomous adversary operating simultaneously across human, non-human, and agentic identity layers.

Adversarial AI Infrastructure: Model Extraction and Vibe Hacking

Beyond state-sponsored espionage, financially motivated threat actors have developed their own agentic playbooks. A methodology now termed "vibe hacking" involves using AI agents to parse massive datasets of exfiltrated information, autonomously identify the most sensitive or legally damaging content, and generate highly tailored extortion demands calibrated to the victim's public financial standing and regulatory exposure, all without deploying traditional ransomware.

Attackers are also targeting enterprise AI infrastructure directly. A significant increase in model extraction attacks has been observed, where threat actors submit hundreds of thousands of prompts to public frontier models to replicate their reasoning capabilities in secondary models that lack safety guardrails entirely. Toolkits such as Xantharox are marketed on dark web forums using these exact methodologies, passing off jailbroken public models as proprietary offensive cyber weapons.

The prompt injection vector is particularly relevant to identity governance. As enterprise AI agents are integrated into internal workflows, including code review, customer

customer service, and data analysis, attackers are embedding hidden malicious commands into the web pages, emails, and documents that these agents are programmed to process. The agent, already trusted and already inside the perimeter, executes the injected instruction with its full set of permissions. Identity governance that does not extend to agentic behavior cannot detect this class of attack.

Section 5: The human cost nobody budgets for

Alert tyranny and the burnout crisis

Behind every security tool and automated detection system sits a human analyst making consequential decisions under chronic stress. The 2026 threat landscape has pushed those conditions past sustainable limits, and the operational consequences are directly measurable.

Modern SIEM environments generate thousands of daily alerts, the overwhelming majority of which are false positives. Analysts cycle through manual triage in a process fundamentally misaligned with a threat environment where attacks are completed in a few minutes. The result is what practitioners now call alert tyranny: a state of chronic exhaustion in which the sheer volume of noise makes genuine signal detection increasingly unreliable.

Eighty-three percent of IT security professionals report that stress and exhaustion have led them or a colleague to make critical errors that directly caused a breach. Eighty-five percent anticipate leaving their current roles due to burnout. 70% of SOC teams report that their personal lives are significantly impacted by work stress. Fifty-five percent report lacking confidence in their ability to prioritize threat responses effectively.

This is not a wellness issue. It is an identity visibility issue. Analysts are overwhelmed precisely because the identity layer across human accounts, non-human identities, and AI agents is not governed uniformly. When every tool produces its own telemetry in isolation, the cognitive load on the human analyst becomes unsustainable.

The paradox of prudence

For security leadership specifically, the threat landscape has introduced a particular challenge. Security professionals are trained to be risk-averse and to delay the

deployment of novel technologies until they are thoroughly proven. In 2026, that instinct has become a liability.

Adversaries do not wait for perfect models or flawless orchestration. GTG-1002 demonstrated that AI agents operating at scale and speed can overwhelm defenders, even while exhibiting hallucinations and tactical inaccuracies. Attackers absorb imperfection through scale. Defenders cannot afford to wait for certainty before deploying autonomous response capabilities.

The risk calculus has shifted. The risk of not deploying AI-driven, identity-aware defense is now measurably greater than the risk of imperfect deployment.

Section 6: What unified identity visibility changes

The architectural gap the threat landscape is exploiting

Every major attack vector in this report shares a common enabling condition: fragmented identity visibility. Security teams operating with separate tools for human identity, non-human identity, and cloud access cannot construct a coherent picture of who or what is acting across their environment at any given moment.

An intrusion that begins with a compromised human credential, pivots through an over-scoped OAuth token, and executes via an AI agent with broad SaaS access will evade detection precisely because no single tool sees the full chain. This is not an edge case. Industry data consistently shows that the overwhelming majority of attacks deliberately span multiple surfaces to exploit exactly this fragmentation.

The solution is not more tools. It is unified visibility across the identity fabric: human, non-human, and agentic, with governance that is continuous, behavioral, and context-aware rather than static and perimeter-bound.

What the Right Architecture Looks Like

For human identities, the response requires real-time behavioral anomaly detection, strict least privilege enforcement, and phishing-resistant MFA that does not rely on employees making correct decisions under AI-generated psychological pressure. A single compromised human identity should not cascade into a full breach. Containment is the design goal.

For non-human identities, the response requires continuous discovery and classification of all service accounts, API keys, OAuth tokens, and machine credentials, with defined ownership, automated rotation, and access scoped strictly to operational need. At a 144:1 NHI-to-human ratio, manual governance is not feasible. Automated, policy-driven lifecycle management is the only viable path.

For AI agents, governance must reflect the fundamental behavioral difference between agentic and traditional non-human identities. Agentic identities require dynamic, context-aware policy evaluation rather than static role assignments. They require behavioral baselines to flag anomalous actions in real time. They require explicit human-in-the-loop approval gates for sensitive or irreversible operations. They require an identity that flows through enterprise SSO and token rotation infrastructure rather than hardcoded credentials that never expire. And they require audit trails that support verifiable data lineage, an emerging requirement under frameworks including the NIST Cyber AI Profile, Texas TRAIGA, and California SB 53.

A practical mental model for security leaders: treat every AI agent as a new hire with no institutional trust. Grant minimum access. Monitor continuously. Escalate consequential actions for human approval.

The same rigor applied to privileged human access must be applied to agentic access, because the blast radius of a compromised agent with broad permissions is equivalent to a compromised privileged insider.

The agentic SOC is the natural evolution of this architecture on the defensive side. Platforms that integrate autonomous security agents to handle Tier 1 and Tier 2 investigative workflows allow human analysts to shift from triage workers to strategic orchestrators, reducing the alert fatigue that is currently driving the retention crisis while simultaneously accelerating detection and response within the 72-minute window.

Conclusion

The identity sprawl nobody is governing

The 72-minute breach is not a benchmark to beat. It is a signal that the underlying architecture of enterprise security needs to change. Every attack pattern documented in this report, from AI-orchestrated espionage to deepfake social engineering to agentic

prompt injection, succeeded not because defenses were absent but because identity was ungoverned. Credentials were over-permissioned. Non-human identities had no owners. AI agents operated with the trust of insiders and the oversight of none.

The path forward is not adding more tools to an already fragmented stack. It is unifying visibility and governance across the full identity fabric: every human and non-human identity, and AI agent, managed under a single, coherent, continuous, behavioral framework built for the speed at which threats now move. Enterprises that make this shift will not just respond faster. They will structurally reduce the attack surface that adversaries have spent the last several years learning to exploit. The window to act is the same one attackers are already moving through.

Unosecur helps security teams unify visibility and governance across human, non-human, and agentic identities, before that gap becomes a breach.

Here is what that looks like in practice:
If your team is struggling to maintain visibility across a rapidly growing estate of non-human identities, service accounts, and AI agents, Unosecur's Unified Identity Fabric maps and monitors them all in one place, continuously, not just at the point of provisioning.

If your analysts are buried in alerts with no unified view of who or what is acting across your environment, Unosecur correlates identity behavior across human, non-human, and agentic layers, so your team sees the full chain of an intrusion, not just fragments.

If your organization is deploying AI agents faster than your IAM frameworks can keep up with, Unosecur brings agentic identities under the same governance model as your privileged human accounts: scoped access, behavioral baselines, and human-in-the-loop controls for consequential actions. The adversaries documented in this report did not succeed because they were technically superior. They succeeded because identity was ungoverned. That is a solvable problem.

Trusted by Security Leaders

Nuno Teodoro, Head of Group Cybersecurity, Solaris

We evaluated top named identity security platforms before choosing Unosecur. What set it apart was the combination and depth of coverage across human and non-human identities, immediate real-time findings from day zero, real-time vs quarterly reviews from classic IGA solutions and an identity timeline built for the audit and regulatory demands we face as a bank. It is our go to platform for identity security and regulatory requirements.

Solaris

+100 more apps & integrations

ISO 27001 Certified · AICPA SOC · GDPR · STAR LEVEL ONE · HIPAA COMPLIANT

unosecur

Unified Identity Security for the Agentic Enterprise