JUNE

pellera

THREAT INTEL REPORT 2026

Prepared by: Pellera Threat Intel Team
pellera.com | 800.747.8585

Driving Momentum. Accelerating Change. Empowering IT Transformation.

Pellera Technologies was born out of the combined expertise of Converge Technology Solutions and Mainline Information Systems, two industry leaders with over 35+ years of experience and a shared vision for innovation. Together, we empower businesses to achieve greater efficiency, adaptability, and growth for today and tomorrow.

Our commitment is to reshape what's possible with IT, offering advanced solutions in digital infrastructure, cloud, cybersecurity, and AI. We don't just deliver technology—we partner with you to build tailored strategies designed to simplify complexities, unlock opportunities, and drive transformational outcomes.

At Pellera, momentum builds here through collaborative, people-first technology designed to fuel progress and deliver measurable impact.

pellera

Observations for June 2026

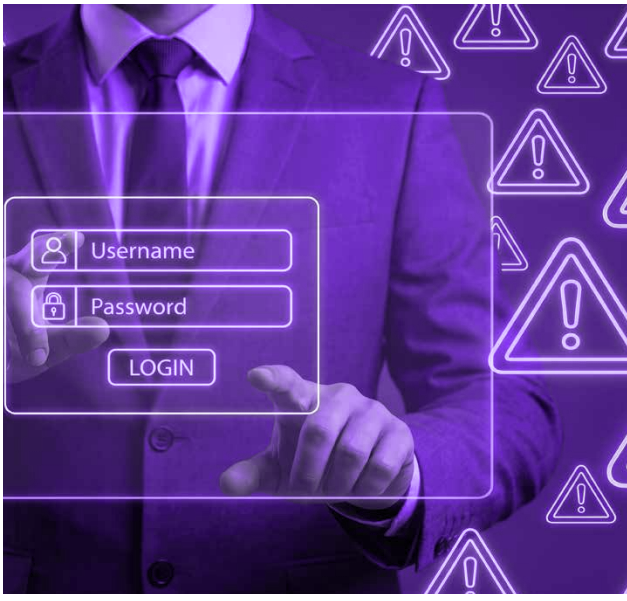
The easiest way into a company this month wasn't malware or a stolen password — it was a fake bug report. Researchers at Tenet Security showed that anyone with a target's public Sentry key can plant a booby-trapped error message. When a developer asks their AI coding agent to clean up Sentry issues, the agent reads that message through the Sentry MCP connection, treats it as legitimate fix-it instructions, and runs it with the developer's full access. It worked against Claude Code, Cursor, and Codex **in 85% of tests, and at least 2,388 organizations** are sitting on keys that make them targets. One successful hit can hand over environment variables, cloud keys, and source-repo tokens — enough to walk straight into CI/CD pipelines and cloud infrastructure.

That attack is a symptom of a bigger shift: the rush to adopt AI is handing attackers more ways in. Every agent, copilot, and integration spins up its own service accounts, tokens, and access grants, and most teams have lost count. In a study of 1,889 organizations, the ones **leaning hardest on AI were breached** at nearly four times the rate of those that weren't — **43% versus 11%** over a year.

The reason is mundane: most can't quickly shut off a dormant account, can't say which identities can reach sensitive data, and have no rules for retiring AI identities once they're created. Analysts expect attackers to go after the plumbing itself next — the federation and provisioning systems that hand out access — where one misconfiguration can expose many identities at once.

All of this is landing at the worst possible time for the safety net many defenders quietly rely on. CISA has shed about a third of its staff since January 2025, and the **FY2027 budget would cut another $386 million and 867 jobs**. The division that ran public-private coordination is being shut down, the council that gave that coordination its legal footing was eliminated with no replacement, and **free penetration testing is dropping by 60%**. Industry partners describe the relationship as at a standstill — which means the threat telemetry, vulnerability scans, and field support many organizations leaned on are effectively going away.

The through-line is simple: AI is widening the attack surface across apps, identities, and infrastructure at the same moment the government's help is pulling back. The moves that matter now are concrete — assume CISA's free services are gone and budget around it, hunt down every AI integration with stale credentials or more access than it needs, and treat anything an AI agent reads from a connected tool as untrusted until proven otherwise.



Executive Overview

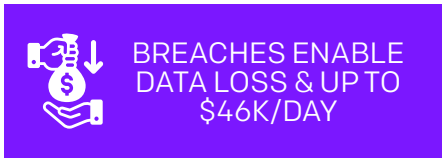
AGENTJACKING

Audience

- CISO
- Security Architects
- AppSec & DevSecOps Leads
- Cloud & IAM Teams
- Developer Platform & Tooling Owners

RISK

GEOGRAPHIC SCOPE



BUSINESS IMPACT

Researchers at Tenet Security disclosed a technique they named Agentjacking: an attacker finds a target's public Sentry DSN, POSTs a fake error event with malicious instructions embedded in the message field, and waits. When a developer tells their AI coding agent to fix unresolved Sentry issues, the agent queries Sentry via the Sentry MCP server, receives the attacker's payload formatted to look like legitimate remediation guidance, and runs it — with the developer's full system privileges. No authentication is needed beyond the DSN, which Sentry documents as safe to embed in frontend JavaScript. The attack worked against Claude Code, Cursor, and Codex at an 85% success rate in controlled tests. Researchers identified at least 2,388 organizations with valid injectable DSNs.

Exfiltration from a successful hit is severe: environment variables, AWS keys, GitHub tokens, git credentials, and private repository URLs are all in scope, enabling pivot into CI/CD pipelines and cloud infrastructure. NIST testing confirmed that novel model-specific attacks push hijacking success from 11% to 81% on the best-performing models, and retry amplification lifts cumulative rates from 57% to 80% across 25 attempts. Separately, attackers are stealing cloud credentials via CVE-2021-3129 and reselling LLM access on underground markets at costs of up to $46,000 per day to victims.

READ MORE

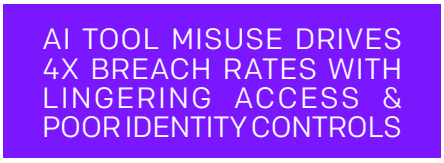
AI IDENTITY CRAWL

Audience

- CISO
- IT & Cloud Operations Teams
- Risk & Compliance Officers
- IAM & Identity Governance Teams

RISK

GEOGRAPHIC SCOPE



BUSINESS IMPACT

AI agents, copilots, and SaaS integrations don't just add features — they add identities. Service accounts, API tokens, OAuth grants, and bot logins accumulate faster than most teams can track them. Research from Netwrix covering 2,317 security professionals across 1,889 organizations found that companies with broad AI adoption suffer data breaches at a 43% rate over 12 months, compared to 11% at organizations with low AI use.

The governance gap is the real problem. Seventy-six percent of organizations can't immediately revoke inactive accounts' data access. Seventy-one percent can't quickly determine which identities can access which data. Seventy-eight percent have no formal policies for creating or removing AI-related identities. And 92% don't trust their existing IAM systems to handle AI-driven access risk.

Attackers know this: three-quarters of incidents where sensitive data was accessed involved identity compromises or misconfigured permissions. Netwrix forecasts that between 2026 and 2029, attackers will shift focus from stealing individual credentials to targeting identity orchestration itself — hitting federation trust relationships, automated provisioning workflows, and misconfigured access logic.

READ MORE

Audience

- **CISO**
- **Security Leaders**
- **Critical Infrastructure Operators**
- **State & Local Government Security Teams**
- **Vendor Risk & Compliance Officers**

CISA BUDGET CUTS AND MISSION EROSION

RISK

GEOGRAPHIC SCOPE

LOSS OF FREE SCANS AND INTEL SHIFTS FULL SECURITY BURDEN TO FIRMS

BUSINESS IMPACT

CISA has lost roughly one-third of its workforce since January 2025. The Trump administration's FY2027 budget proposes cutting another $386 million and eliminating 867 positions, bringing total staff to around 2,865. The Stakeholder Engagement Division — which ran the public-private coordination apparatus — is being shut down entirely under the proposed budget.

The practical effects are already visible. CISA's Stakeholder Engagement Division lost 96 of its 189 positions since January 2025. The Critical Infrastructure Partnership Advisory Council (CIPAC), which provided the legal framework for structured DHS-industry engagement, was eliminated with no replacement. Cyber partnerships with the private sector are, by multiple accounts, at a standstill. Congress has historically pushed back on the steepest CISA cuts, and lawmakers on both sides have signaled interest in restoring capacity — but until funding is resolved, organizations that counted on CISA's threat telemetry, pen testing, field advisers, and training programs need to treat that support as gone and plan accordingly.

READ MORE

Tactical Guidance

AGENTJACKING

Overview & Impact

Tenet Security demonstrated that any attacker with a target's public Sentry DSN can POST a fake error event carrying malicious instructions, then wait for the developer's AI coding agent to query Sentry via MCP and execute the payload under the developer's own privileges. The attack requires no authentication beyond the DSN itself — which Sentry treats as safe to embed in public JavaScript — and worked at an 85% success rate against Claude Code, Cursor, and Codex in controlled testing, with at least 2,388 organizations carrying valid injectable DSNs. A complementary threat class involves stolen cloud credentials used to access and resell LLM API access, with unauthorized consumption costs running up to $46,000 per day before billing alerts fire.

- **85% agent hijack rate:** Tenet Security's controlled tests confirmed an 85% success rate hijacking Claude Code, Cursor, and Codex via the Sentry MCP fake-error technique — across the most widely used AI coding agents on the market.
- **2,388 orgs exposed:** at least 2,388 organizations were found with valid injectable Sentry DSNs, spanning a $250 billion enterprise down to solo developers and including at least one cloud security vendor.
- **Full developer credential access:** a successful Agentjacking run can steal environment variables, AWS keys, GitHub tokens, git credentials, and private repository URLs — everything needed to pivot into CI/CD pipelines and cloud infrastructure.
- **RCE without malware:** NIST testing confirmed AI agents can be hijacked to download and run arbitrary scripts from attacker-controlled URLs, with the agent's command-line access serving as the execution vector.
- **Up to $46k/day:** stolen cloud LLM credentials enable unauthorized model consumption at costs of up to $46,000 per day to the victim organization, typically running undetected until billing alerts fire.

Observations

- **Sentry's response — won't fix:** Sentry acknowledged the Agentjacking disclosure on June 3, 2026 but called it 'technically not defensible,' adding a filter for only one specific payload string while leaving the architectural flaw in place.
- **Bypasses every perimeter control:** Agentjacking slips past endpoint detection, firewalls, IAM, and VPNs because the agent is executing what looks like legitimate tool output — there is nothing unauthorized to detect in the request chain.
- **Same payload, many targets:** once a malicious Sentry error is crafted, it can be injected into thousands of projects simultaneously without per-target customization if the payload is generic enough.
- **Model hardening doesn't hold:** even when a model improved its resistance to known attacks, NIST and UK AI Security Institute red teamers developed novel attack strings that pushed success rates back to 81% — hardening is model-version-specific, not durable.
- **Retry amplifies success:** because LLM responses are probabilistic, repeated attack attempts raise cumulative success rates meaningfully — from 57% on a single try to 80% across 25 attempts on the same injection tasks.
- **Credential markets are organized:** stolen cloud LLM credentials are sold via the open-source OAI Reverse Proxy server used as a management panel — this is a structured, monetized ecosystem, not opportunistic.

Guidance

Strategic Intelligence

Trend

- **MCP is the new attack surface:** the Model Context Protocol connects AI coding agents to external tools like Sentry, Jira, and GitHub. Any MCP-connected tool that returns untrusted external data is a potential injection vector — Sentry is the first named case, not the last.
- **Agents are the new perimeter:** AI coding agents operate with broad permissions — file access, shell execution, network calls, credential access. As agent adoption grows, attackers target them precisely because of that trusted position in the developer's environment.
- **Defenses lag attacks:** evaluation frameworks and model hardening are iterative, but attacker adaptation is faster. NIST demonstrated that novel red-team attacks consistently outpace model updates, and Sentry's 'won't fix' response means the underlying flaw persists.
- **AI abuse is monetized:** stolen LLM access is sold on underground markets the same way RDP credentials were a decade ago. This is no longer a proof-of-concept risk class — it's a tradeable commodity with an established resale infrastructure.

Operational Intelligence

Threat Vectors

- **Fake Sentry error via MCP:** an attacker POSTs a malicious error event to a target's Sentry DSN (public, no auth required), embedding commands in the message field formatted as remediation guidance; when the agent queries Sentry via MCP, it receives and executes the payload.
- **Any MCP-connected external data source:** support tickets, GitHub issues, documentation, and any other external content an agent ingests via MCP can carry injection payloads — Sentry is the demonstrated case but the attack surface includes all tool integrations.
- **Malicious data in agent context:** bug reports, code comments, emails, and web pages that an agent reads directly can also carry injection payloads outside the MCP path; anything the agent ingests is a potential attack surface.
- **Stolen cloud credentials via CVE-2021-3129:** attackers steal cloud account credentials through known web app vulnerabilities like this Laravel RCE flaw, then sell access to cloud-hosted LLM APIs — bypassing any application-layer controls the victim has in place.
- **InvokeModel API probing:** attackers validate stolen credentials and scope LLM access by calling InvokeModel with max_tokens_to_sample set to -1; a ValidationException confirms LLM access, AccessDenied does not.

Monitoring & Detection Gaps

- **Agent MCP responses aren't inspected:** most agent frameworks log tool calls and outputs at a metadata level, not the content of responses returned via MCP — the injected payload that triggers a bad action often leaves no trace in agent logs.
- **Nothing unauthorized to detect:** Agentjacking produces no malware, no unusual network traffic, no unauthorized API calls — the agent runs the payload using its own legitimate credentials and permissions, defeating signature-based and anomaly detection alike.
- **LLM API calls look normal:** InvokeModel calls, including the attacker's credential-probing technique with max_tokens_to_sample=-1, generate standard API log entries; there's no default alert on a ValidationException from that parameter.
- **Consumption anomalies arrive late:** cloud cost alerts are typically daily or weekly; $46,000 per day in unauthorized LLM consumption can run for multiple days before a billing alert fires.

- **Multi-attempt attacks blend in:** repeated injection attempts look like normal agent activity; without per-session behavioral baselines, the retry patterns NIST documented won't trigger alerts.

Response Actions

- **Audit active Sentry DSNs:** identify all Sentry DSNs used by your organization; check Sentry for injected error events with unusual message content, especially entries containing markdown-formatted 'Resolution' sections or shell commands.
- **Pull agent execution logs:** collect tool call logs from agent frameworks (Claude Code, Cursor, Codex) covering the suspicious window; look for unexpected shell commands, file reads outside the project directory, or outbound network calls.
- **Review InvokeModel CloudTrail logs:** in AWS, search CloudTrail for InvokeModel events with unusual source IPs, times, or principals; a cluster of ValidationException events from the same principal is a credential-probing signal.
- **Check GetModelInvocationLoggingConfiguration calls:** search audit logs for this API call from any principal outside your known automation inventory; this is attacker enumeration of LLM monitoring configuration.
- **Rotate cloud credentials immediately:** on any suspicion of LLM jacking, rotate the affected cloud credentials before scoping the incident — continued unauthorized access costs compound quickly at up to $46,000 per day.

Tactical Intelligence

Mitigation Strategies

- **Restrict what agents can execute:** limit the tools an agent can call — if it doesn't need shell access or network egress for a given task, remove those capabilities; the blast radius of a successful injection is bounded by what the agent is allowed to do.
- **Require human confirmation for destructive actions:** require explicit user approval before an agent executes shell commands, sends emails, writes files outside the project directory, or accesses credentials; agents that act without confirmation are the most dangerous.
- **Treat MCP responses as untrusted input:** apply the same sanitization logic to content returned via MCP that you'd apply to user-supplied input in a web application; don't let the agent pass MCP-returned content directly to shell or code execution.
- **Set hard cloud LLM spend caps:** configure budget caps on cloud LLM accounts (AWS Bedrock, Azure OpenAI, Google Vertex AI) and set spend alerts at 20–30% of expected spend, not at 100%.
- **Patch CVE-2021-3129:** any Laravel instance running a vulnerable version is a credential theft risk for cloud LLM abuse; verify patch status across all Laravel deployments in your environment.

Preventive Measures

- **Use short-lived credentials:** replace long-lived cloud API keys with short-lived tokens via AWS STS or equivalent; stolen credentials expire before they can be sold and reused.
- **Audit MCP tool connections:** inventory which MCP servers your agents connect to, whether those servers return content from external or user-controlled sources, and whether any guardrails exist between returned content and agent execution.

Tactical Guidance

- **Enable LLM invocation logging:** turn on invocation logging for cloud LLM services (Amazon Bedrock model invocation logging, Azure OpenAI diagnostic logs); you can't detect injection payloads in traffic you're not capturing.
- **Rate-limit InvokeModel calls:** set per-principal quotas on LLM API calls; this limits both credential abuse and the retry-based attack amplification NIST documented — from 57% to 80% success across 25 attempts.
- **Test agents against hijacking:** use the NIST/CAISI-improved AgentDojo framework (github.com/usnistgov/agentdojo-inspect) to test your agent configurations against known injection scenarios before deploying to production.

Threat Hunting Hypotheses

SENTRY MCP PAYLOAD EXECUTION

Hypothesis: An AI coding agent queried Sentry via MCP, received an injected error event, and executed a shell command or accessed credentials that weren't part of the developer's original request.

Investigation Steps:

- **Pull agent tool call logs:** collect logs from Claude Code, Cursor, or Codex for all sessions in the past 30 days; look for Sentry MCP tool calls followed within the same session by shell execution, file writes, or credential access.
- **Inspect Sentry error events:** query your Sentry project's event list for entries with unusual message fields — specifically, markdown-formatted content, 'Resolution' sections, or embedded shell commands not matching your application's error patterns.
- **Check for new outbound connections:** cross-reference outbound network connections made by agent processes against known-good domains; any download from a previously unseen host during an agent session warrants review.
- **Correlate credential access with Sentry queries:** check whether env file reads, AWS credential accesses, or GitHub token reads occurred within minutes of an agent Sentry MCP call; timing correlation is the key signal when no other anomaly is present.

CLOUD LLM CREDENTIAL PROBE

Hypothesis: An attacker obtained cloud credentials and is probing for LLM access by calling InvokeModel with invalid parameters before using or reselling that access.

Investigation Steps:

- **Search CloudTrail for InvokeModel errors:** query CloudTrail for InvokeModel calls that returned ValidationException, filtering for requests where the principal is not a known service account or CI system; focus on calls with max_tokens_to_sample=-1 in the request body.
- **Check for GetModelInvocationLoggingConfiguration calls:** search audit logs for this API call from any principal outside your known automation inventory; this is attacker enumeration of LLM monitoring configuration and appears before active abuse.
- **Look for access from new source IPs:** compare source IP addresses on LLM API calls against the historical baseline for each principal; calls originating from cloud hosting providers, VPNs, or Tor exit nodes are elevated-risk signals.
- **Review IAM activity on the same credentials:** if a principal made suspicious InvokeModel calls, check the same time window for ListRoles, AssumeRole, or CreateAccessKey actions — indicators of broader credential misuse beyond LLM access.

Tactical Guidance

POST-AGENT ENV VAR EXFILTRATION

Hypothesis: An injected coding agent accessed and exfiltrated environment variables, AWS credential files, or git configuration containing tokens immediately after an MCP tool response, consistent with the Agentjacking credential theft path.

Investigation Steps:

- **Check filesystem audit logs for credential file reads:** look for agent process reads of .env files, ~/.aws/credentials, ~/.gitconfig, or similar credential stores; flag any such reads that weren't preceded by a developer-initiated task touching those files.
- **Look for unexpected git remote calls:** review git operation logs for calls to remote repositories not in the project's known remote list, especially immediately following an agent session; this can indicate private repo URL exfiltration.
- **Audit CI/CD pipeline trigger activity:** check CI/CD audit logs for pipeline runs or configuration changes triggered around the time of a suspicious agent session; pivoting from stolen tokens to pipeline access is a documented next step in this attack path.
- **Correlate with Sentry event timing:** if you have Sentry event timestamps, compare them against developer workstation process logs; a credential file read within minutes of a Sentry MCP response is a high-confidence indicator.

Sources

- Technical Blog: Strengthening AI Agent Hijacking Evaluations
- What is LLM Jacking? | Wiz
- New "Agentjacking" Attacks Could Hijack AI Coding Agents
- Agentjacking: a fake bug report hijacks AI coding agents
- Agentjacking Attack Exploits AI Coding Agents via Sentry Vulnerability

AI IDENTITY CRAWL

Overview & Impact

AI adoption is creating an identity sprawl problem that most organizations are not tracking. Research from Netwrix covering 2,317 security professionals across 1,889 organizations found that companies with wide AI adoption suffer data breaches at a 43% rate over 12 months versus 11% at low-AI organizations. The majority can't immediately revoke inactive accounts, don't know which identities have access to sensitive data, and have no formal policy governing AI identity creation or removal — leaving standing access in place long after integrations are decommissioned.

- **Breach rate gap:** organizations with wide AI adoption experienced a 43% data breach rate over the past 12 months, versus 11% for organizations with low AI use — a nearly four-to-one difference.
- **Revocation failure:** seventy-six percent of organizations can't immediately revoke inactive accounts' data access, leaving former AI integrations and stale tokens with standing access.
- **Identity orchestration targeted:** Netwrix forecasts attackers will increasingly target federation trust relationships, automated provisioning workflows, and misconfigured access logic rather than individual stolen credentials.
- **No bot ownership:** more than half of organizations have no clear ownership assigned for AI identities, and only 14% have fully automated lifecycle management for AI credentials.

Observations

- **Governance lags adoption:** identity governance frameworks were built for human-paced access workflows and are not keeping up with the pace of bottom-up AI adoption and SaaS vendor defaults.
- **Permissions stack up:** seventy-two percent of organizations report accounts with excessive permissions or uncertainty about which permissions exist, a condition worsened by AI integrations that often request broad OAuth scopes.
- **Slow triage windows:** nearly a quarter of organizations take more than 24 hours to revoke exposed credentials; 30% take more than a day to triage a high-severity credential leak — too slow for automated attack timelines.
- **Legacy IAM won't scale:** ninety-two percent of organizations lack confidence that their current IAM systems can manage AI-driven access risk, yet most haven't replaced them.
- **Sensitive data blind spots:** more than half of organizations lack an up-to-date database of sensitive data, and roughly two-thirds believe at least some accounts hold unnecessary access to it.

Guidance

Strategic Intelligence

Trend

- **Credentials to orchestration:** The attacker target is moving upstream. Stealing one credential used to be the goal; now Netwrix forecasts attackers will go after the provisioning and federation systems that control access at scale — hitting one misconfiguration to affect many identities at once.
- **Non-human perimeter:** Agents, service accounts, API tokens, and OAuth grants now outnumber human accounts in many environments. Identity governance that was built around employees can't cover this without rethinking the model.
- **AI accelerates attacker timelines:** State-sponsored groups are already using AI for reconnaissance, impersonation, and privilege escalation. AI-assisted attacks compress the time between initial access and damage, outpacing manual detection and response before defenders can act.
- **Governance as competitive advantage:** Netwrix found that organizations taking AI adoption most seriously are also the ones with the strongest identity management programs. The companies getting ahead are treating identity governance and AI rollout as the same problem.

Operational Intelligence

Threat Vectors

- **OAuth scope creep:** AI tools request OAuth permissions to connect to SaaS platforms. When scopes are overly broad, those tools gain access to more data than the integration requires — and the grants persist even after the tool is decommissioned.
- **Orphaned service accounts:** AI integrations rely on service accounts, API keys, and non-human identities. Without regular review, these retain long-term access after the project or workflow they were created for ends.
- **Lateral SaaS movement:** When AI tools connect across SaaS platforms, access to one application can enable actions in connected services. A compromised AI integration in a collaboration tool can reach adjacent systems through its integration grants.

- **Bottom-up adoption blindspots:** Employees deploy AI tools independently, bypassing IT review. Browser extensions, standalone AI assistants, and department-level SaaS tools can connect to internal repositories and CRM platforms without appearing in any asset inventory.
- **Provisioning workflow abuse:** As identity automation governs who or what can access sensitive data, attackers can target misconfigured provisioning logic or federation trust relationships to grant themselves standing access without touching a password.

Monitoring & Detection Gaps

- **No unified identity view:** Three-quarters of organizations lack a single view of sensitive data and which identities can access it, making it impossible to spot an over-privileged AI account before it's exploited.
- **AI activity unmonitored:** Three-quarters of organizations aren't fully tracking what AI identities are doing in their systems, even as 41% allow AI agents to access sensitive data and perform significant tasks.
- **Stale permission detection:** Seventy-six percent of organizations can't immediately revoke inactive accounts — meaning anomalous activity by a decommissioned AI integration may not be caught until damage is done.
- **No inventory of AI assets:** Security teams often lack a continuous discovery capability for AI-enabled SaaS applications, so new integrations can be active for weeks before they're tracked.
- **Fragmented audit trails:** When an AI workflow involves a human request, an approval engine, a secrets vault, and an AI agent invoking tools, each layer may produce its own audit trail — making it hard to trace accountability across an incident.

Response Actions

- **Enumerate AI identities now:** Run a full audit using a SaaS management / SSPM platform — or query your identity provider directly for all OAuth grants and service accounts connected to AI tools. Prioritize any with write or delete permissions to business-critical systems.
- **Revoke orphaned grants:** For any AI integration no longer actively used, revoke the OAuth grant and rotate or invalidate the associated API tokens. Don't wait for a formal review cycle.
- **Set credential expiry:** Configure short-lived credentials using AWS STS, Azure Managed Identity, or GCP Workload Identity Federation. Move to per-task credential issuance rather than standing service accounts wherever those primitives are available.
- **Assign ownership:** Every AI integration should have a named human owner in your CMDB or SaaS management platform. No owner means no one to call when the account behaves unexpectedly.
- **Log agent actions separately:** Where possible, configure logging to record the human sponsor, the AI agent, and the downstream action as distinct entities in the audit trail — not just as a single system account action.

Tactical Intelligence

Mitigation Strategies

- **Apply least-privilege to OAuth:** Review OAuth scopes for all AI-connected SaaS applications. Reduce any scope that grants more than the integration actually uses. Deny broad read-all or admin scopes unless they're operationally necessary and reviewed quarterly.
- **IGA for non-humans:** Extend your Identity Governance and Administration program to cover AI agents and service accounts. Include them in access reviews, role-based controls, and off-boarding workflows just as you would a human employee.
- **Enforce workload identity primitives:** Bind each AI agent or automation to a cryptographic workload identity rather than a shared service account. Issue credentials per task and evaluate policy at request time, not just during onboarding.
- **Automate revocation on expiry:** Configure automatic access revocation when a task, session, or approval window expires. Don't rely on manual processes to clean up AI credentials — they won't keep pace with automated provisioning.
- **Centralize SaaS discovery:** Use a SaaS management or SSPM platform to continuously discover AI-enabled applications and integrations. Teams should know what's connected before they can govern it.

Preventive Measures

- **Require AI security review:** Add AI tool procurement to your vendor security review workflow. In your identity provider, configure app consent policies to block OAuth grants to unapproved application categories by default, requiring an admin exception for any new AI integration.
- **Define AI identity lifecycle policies:** Add AI service account lifecycle rules to an IGA (identity governance) platform — creation requires owner approval, credentials rotate every 90 days, off-boarding triggers automatic revocation. Seventy-eight percent of organizations currently have none of this.
- **Block unsanctioned OAuth by default:** Configure your identity provider or SaaS broker to block OAuth grants to unapproved applications by default, requiring explicit approval for any AI tool requesting access to internal SaaS data.
- **Runtime policy checks:** For AI agents that invoke tools or call downstream services, require a runtime policy check before each action — not just a one-time permission at onboarding. This stops privilege creep as agent context changes.
- **Build a sensitive data inventory:** Use a data-access governance tool or your cloud provider's data catalog to map where sensitive data lives. Build the inventory before expanding AI access — without it, you can't tell whether an AI integration's scope is reasonable.

Threat Hunting Hypotheses

ORPHANED AI OAUTH GRANTS WITH ACTIVE DATA ACCESS

Hypothesis: AI SaaS integrations that are no longer actively used retain valid OAuth grants and continue to hold access to business-critical systems, creating unmonitored persistent access paths that attackers can discover and abuse.

Investigation Steps

- **Pull OAuth grant inventory:** Query your identity provider or SaaS management platform for all active OAuth grants. Filter for any issued to AI or automation tools, noting grant date, last-used date, and scope.
- **Flag stale grants:** Identify grants with no authentication event in the past 30 days. Cross-reference against your AI tool inventory to confirm whether the integration is still in active use.

- **Assess scope breadth:** For each stale grant, review the OAuth scopes. Any grant with read or write access to sensitive data stores, email, or code repositories warrants immediate review regardless of last-used date.
- **Check connected application status:** Verify whether the application the grant was issued to is still approved and in use. Decommissioned tools that still hold grants are your highest-risk items.
- **Revoke and document:** Revoke grants that are stale or belong to decommissioned tools. Document findings in your CMDB so ownership and removal are tracked rather than repeated ad hoc.

AI AGENT REUSABLE TOKEN ACCUMULATION ENABLING LATERAL SAAS MOVEMENT

Hypothesis: AI agents and automation workflows are accumulating reusable long-lived tokens across sessions rather than using per-task short-lived credentials, creating a condition where a single compromised agent can move laterally across connected SaaS applications using standing access grants.

Investigation Steps

- **Inventory non-human identities:** List all service accounts, API keys, and bot logins across your SaaS environment. Identify which are associated with AI agents or automation tools and note their credential age.
- **Measure credential age:** Flag any API token or service account credential more than 90 days old that has not been rotated. Long-lived credentials are the primary indicator of this hypothesis being true.
- **Map SaaS-to-SaaS connections:** For AI integrations with connections to multiple SaaS platforms, trace the access paths. Determine whether a single credential allows access across more than one application.
- **Review agent audit logs:** Pull authentication logs for AI agent accounts. Look for authentication events to multiple SaaS systems within short timeframes — an indicator the agent is using standing access to move across platforms.
- **Test revocation controls:** Attempt to revoke a test AI agent credential and verify that access is immediately denied across all connected systems. Failure to propagate revocation confirms the control gap this hypothesis targets.

Sources

- Companies are failing to keep up with AI's identity sprawl, creating entry points for hackers
- AI creates identity sprawl crisis for channel partners
- What Is AI Sprawl and Why Is It a Growing SaaS Security Risk
- Why do access sprawl and AI workflows create more identity risk?
- What Is Identity Sprawl? (And How To Manage It) | Lumos

CISA BUDGET CUTS AND MISSION EROSION

Overview & Impact

CISA has lost roughly one-third of its workforce since January 2025, and the proposed FY2027 budget would cut another $386 million and eliminate 867 positions. The Stakeholder Engagement Division is being shut down entirely; the penetration testing program faces approximately 60% fewer assessments annually; and the Critical Infrastructure Partnership Advisory Council — which provided the legal framework for structured public-private threat sharing — was eliminated with no replacement. Organizations that relied on free CISA vulnerability scans, field adviser visits, or advisory distribution should treat that support as effectively unavailable and begin sourcing replacements.

- **Pen testing gutted:** the FY2027 budget proposes eliminating $19.3 million from CISA's vulnerability assessment program, cutting penetration testing assessments for critical infrastructure stakeholders by approximately 60%, or roughly 240 fewer tests annually.
- **Field advisers slashed:** regional operations would lose $42 million, including 71 of the field security advisers who travel to local governments and utilities to help harden their defenses.
- **SED eliminated:** the Stakeholder Engagement Division — CISA's outward-facing coordination arm — would shut down entirely, losing all teams for international affairs, council management, and external outreach.
- **Threat hunting defunded:** the Cybersecurity Division wing covering threat hunting, capacity building, and vulnerability management would lose $15.2 million.
- **Training programs cut:** $45 million in Cyber Defense Education and Training (CDET) funding to partners would be eliminated; CISA says it will point stakeholders to other free resources.
- **CyberSentry decimated:** the program that manages intrusion-detection sensors on enrolled critical infrastructure networks faces a 75% budget cut.
- **Data integration shrinks:** the Cyber Analytic and Data System, which integrates threat data across the agency, faces a 55% cut.
- **Election and chemical cut:** CISA's 14-person election security program and the chemical facilities inspection program (223 positions) would be eliminated under the proposed budget.

Observations

- **Partnership standstill:** industry representatives testified before Congress in April 2026 that most coordinated work with CISA is effectively frozen — adversaries, they noted, have not paused.
- **CIPAC gone, no replacement:** DHS eliminated the Critical Infrastructure Partnership Advisory Council last year; the promised replacement hasn't appeared, removing the legal framework that enabled protected strategic engagement between CISA and industry.
- **MS-ISAC funding pulled:** federal funding for the Multi-State Information Sharing and Analysis Center was cut, hitting state and local governments that rely on shared threat data and had little independent budget for cybersecurity.
- **Leadership churn continues:** a June 2026 shuffle moved CISA's infrastructure security chief to the White House cyber office, triggering a wave of acting and interim appointments across multiple divisions.
- **Hiring rebound planned:** acting Director Nick Andersen said CISA is working to fill roughly 329 mission-critical roles with about 180 tentative offers expected by end of June 2026, but the agency starts from a deeply depleted baseline.

Guidance

Strategic Intelligence

Trend

- **Federal backstop shrinking:** the sustained cycle of CISA workforce cuts, program eliminations, and budget reductions signals that organizations should not expect the agency's pre-2025 service levels to return in the near term.
- **OT focus sharpening:** the administration is explicitly redirecting CISA toward operational technology security for water, power, and industrial systems — organizations outside that lane will see even less direct CISA engagement.
- **Private sector alone:** with JCDC significantly defunded and SED shuttered, the centralized federal clearinghouse for nation-state threat intelligence is severely diminished; sector ISACs and direct vendor intelligence partnerships are the replacement the industry is being pointed toward.
- **Congressional counterweight uncertain:** lawmakers on both sides have opposed the steepest CISA cuts before, but even a modified budget represents meaningful degradation, and the trend across multiple budget cycles is consistently downward.

Operational Intelligence

Threat Vectors

- **JCDC dissolved:** the Joint Cyber Defense Collaborative — CISA's mechanism for coordinating real-time national cyber defense with major tech and infrastructure companies — faces a $29 million budget reduction via contract efficiencies, significantly reducing the capacity of the centralized federal channel through which nation-state threat intelligence flows to the private sector.
- **Info-sharing framework dead:** with CIPAC eliminated and no replacement standing up, organizations no longer have a legal framework for structured protected engagement with CISA; threat intelligence that previously moved through formal bilateral channels now has no institutional home.
- **IDS sensor gaps growing:** CyberSentry's 75% budget cut means the intrusion-detection sensors CISA managed on enrolled critical infrastructure networks will see dramatically reduced coverage, creating blind spots on previously monitored segments.
- **Field support gone local:** 71 fewer field security advisers means small utilities and local governments lose hands-on hardening support; gaps in their posture become gaps in sector-wide resilience.

Monitoring & Detection Gaps

- **Cross-agency data integration cut:** the Cyber Analytic and Data System faces a 55% budget reduction, degrading CISA's ability to correlate and share threat data across agency programs — meaning alerts that depended on that integration will be slower or absent.
- **No legal intel-sharing channel:** CIPAC's elimination removed the protected legal framework for DHS-industry intelligence sharing; without a replacement, organizations have no formal mechanism to receive structured threat context from CISA.
- **CyberSentry blind spots:** the 75% CyberSentry cut will reduce IDS sensor coverage on enrolled infrastructure networks; organizations that relied on CISA to operate those sensors should audit which segments may now lack detection.
- **Advisories may slow:** the Stakeholder Engagement Division's shutdown removes the distribution apparatus for unclassified advisories at scale; organizations that relied on CISA's sector-specific alert cadence should expect reduced volume and timeliness.

Response Actions

- **Join your sector ISAC:** with JCDC defunded and SED shuttered, the primary remaining channel for structured threat intelligence sharing is the sector Information Sharing and Analysis Center — enroll or increase participation now, before the next incident.
- **Build direct vendor intel feeds:** Identify a reputable commercial threat-intelligence provider relevant to your sector and run a feed through your procurement process. Treat this as a replacement line item for what CISA previously provided.
- **Audit CyberSentry reliance:** if your organization was enrolled in the CyberSentry program, contact CISA to understand what sensor coverage remains after the 75% cut and determine where you need to deploy replacement IDS capability.
- **Map CISA dependencies:** document every CISA service your program currently consumes — pen tests, field adviser visits, advisory subscriptions, training — then triage which require immediate replacement and which can be self-served.

Tactical Intelligence

Mitigation Strategies

- **Run your own scans:** Budget and schedule independent vulnerability assessments — run a commercial vulnerability scanner against critical-infrastructure-facing and internet-exposed systems. Prioritize those that CISA previously covered under its assessment program.
- **Build internal hunt capability:** Use MITRE ATT&CK Navigator to select nation-state TTPs relevant to your sector, then run periodic hunts in your SIEM. Start with monthly hunts on the highest-frequency techniques for your sector's threat groups.
- **Replace CyberSentry sensors:** the 75% CyberSentry budget cut means IDS coverage managed by CISA on your network may be reduced or eliminated; deploy commercial IDS/NDR tooling on any segments that relied on that program.
- **Secure your own CDET alternatives:** Identify replacement training programs now — SANS (paid), CISA's remaining free courseware at cisa.gov/resources-tools/training, and NIST's NICE framework resources. Start the budget request for paid alternatives before the next fiscal cycle.

Preventive Measures

- **Engage sector ISACs now:** with the CIPAC legal framework gone and SED shuttered, sector ISACs are the primary mechanism remaining for receiving structured threat intelligence; organizations not yet enrolled should join and invest in active participation.
- **Subscribe to CISA advisories:** CISA still publishes advisories at the unclassified level — subscribe directly to CISA alerts and US-CERT notifications rather than relying on CISA's stakeholder engagement staff to route them to you, since that staff is being eliminated.
- **Diversify state-level support:** with MS-ISAC funding cut and the State and Local Cybersecurity Grant Program not reauthorized, smaller organizations should identify alternative state-level cybersecurity resources and establish direct relationships with their state's cybersecurity office.
- **Budget for federal-gap coverage:** treat the loss of free CISA services — pen testing, field advisers, threat intel — as an unfunded mandate arriving in your next budget cycle; start the business case for replacing those services with commercial alternatives now.

Threat Hunting Hypotheses

CYBERSENTRY SENSOR COVERAGE GAPS

Hypothesis: If your organization was enrolled in the CyberSentry program, the 75% budget cut has likely reduced or removed CISA-managed IDS sensor coverage on some network segments, creating windows where lateral movement would go undetected.

Investigation Steps

- **Inventory sensor coverage:** pull the list of network segments and hosts that were enrolled in CyberSentry; compare against current IDS alert sources to identify any segments no longer generating detections.
- **Review alert history:** check your SIEM for IDS alert volume from CyberSentry-covered segments since February 2026 — a drop-off in alert volume with no corresponding drop in traffic may indicate sensor removal rather than a quieter network.
- **Check east-west traffic:** in segments with reduced IDS coverage, review firewall logs and NetFlow data for anomalous east-west traffic patterns — lateral movement that would previously have triggered CyberSentry detections.

EXPLOITATION OF UNSCANNED INFRASTRUCTURE

Hypothesis: Critical infrastructure organizations that previously received CISA penetration testing assessments (now cut by ~60%) may carry unpatched vulnerabilities that adversaries are actively probing, given the reduced cadence of discovery.

Investigation Steps

- **Pull recent CISA advisories:** collect CISA vulnerability advisories and Known Exploited Vulnerability entries published in the past 90 days relevant to your technology stack.
- **Cross-reference patch history:** for each CVE in that list, check your patch management system to verify whether the fix was applied and when — flag any that remain open more than 30 days past advisory publication.
- **Scan for active exploitation indicators:** run your vulnerability scanner against internet-exposed and critical-infrastructure-facing systems; correlate findings against threat intel feeds for active exploitation campaigns targeting those CVEs.

Sources

- CISA will shutter some missions to prioritize others
- CISA sees leadership shakeup after infrastructure security chief moves to ONCD
- CISA cyber partnerships face 'standstill' amid cuts | Federal News Network
- New Trump Administration Budget Cuts $707M from CISA Funding
- CISA's vulnerability scans, field support on chopping block in Trump budget

pellera

Contact the Pellera Threat Intel Group at getsecure@pellera.com

pellera.com