INTELLIGENCE BRIEFING
Security Command Center
TLP:CLEAR
2026-04-30 14:09 UTC

Shadow AI Inventory Gap Exposes Enterprises: Real Deployments Exceed Self-Reported Counts by 3x or More

GOVERNANCE | MEDIUM | CVSS 5.0

SCC Item ID	SCC-GOV-2026-0023
Type	Governance
Severity	MEDIUM
CVSS Base Score	5.0
Affected Products	Enterprise SaaS environments, AI agents, copilots, and browser extensions (vendor-agnostic); CrowdStrike Falcon Platform (Shadow AI Visibility Service, Falcon AIDR)
Discovery Source	Rss:T1 Threatintel

Executive Summary

CrowdStrike field data shows enterprises have three or more times as many active AI deployments as their own inventories reflect. Untracked AI agents, copilots, and browser extensions operate under inherited user permissions with no security oversight, creating blind spots across endpoints, SaaS platforms, and cloud storage. The core business risk is ungoverned autonomous action: these agents can read, move, or exfiltrate sensitive data with no audit trail and no human in the loop.

Technical Analysis

This is a structural governance and visibility gap, not a patchable vulnerability. CrowdStrike field engagements document systematic undercounting of AI deployments across enterprise SaaS, endpoint, and cloud environments. Affected weakness patterns: CWE-285 (AI agents operating under excessive inherited user-level permissions with no least-privilege enforcement), CWE-200 (sensitive data submitted to external LLMs at the prompt layer, bypassing DLP controls), and CWE-693 (shadow AI processes bypassing web filtering and policy controls). Relevant MITRE ATT&CK techniques include T1078 (Valid Accounts, agents inherit user token permissions), T1567 (Exfiltration Over Web Service, data sent to external LLM APIs), T1530 (Data from Cloud Storage, agent access to connected cloud data), T1059 (Command and Scripting Interpreter, agent-executed commands on connected systems), T1526 (Cloud Service Discovery), T1213 (Data from Information Repositories), T1190 (Exploitation of Remote Services), and T1195 (Supply Chain Compromise via MCP-connected and AI-integrated dev tools). CrowdStrike's remediation posture: Falcon AIDR (GA December 2025) for AI-specific detection and response; Shadow AI Visibility Service (launching April 2026) for inventory and footprint mapping; unified data protection for GenAI prompt-layer DLP. No CVE applies. Source quality is

vendor-sourced (T3); independent third-party corroboration is not available in this data set.

Action Checklist

1. Step 1: Inventory, Run an AI asset discovery sweep across endpoints, SaaS integrations, browser extensions, and cloud-connected services. Do not rely on self-reported counts or IT procurement records. CrowdStrike's field data indicates those undercounts by 3x or more. Use endpoint telemetry and network egress logs to surface undeclared AI processes.
2. Step 2: Detection, Query endpoint logs for processes or browser extensions making outbound connections to known LLM API endpoints (api.openai.com, api.anthropic.com, generativelanguage.googleapis.com, and equivalents). Flag any agent operating under a standard user token accessing cloud storage (T1530) or submitting data to external web services (T1567). Review SaaS OAuth grant lists for AI applications with broad permission scopes.
3. Step 3: Eradication, Revoke OAuth grants for unvetted AI applications. Apply least-privilege to any AI agent or copilot that inherited broad user-level permissions. Block egress to unapproved LLM API endpoints at the proxy or firewall layer. Remove unsanctioned AI browser extensions via endpoint management policy.
4. Step 4: Recovery, Validate that your AI application inventory now reflects actual deployment counts via telemetry, not self-reporting. Confirm DLP controls cover prompt-layer traffic to external LLM APIs. Monitor OAuth grant lists and browser extension inventories on a recurring schedule (weekly minimum during remediation phase).
5. Step 5: Post-Incident, Establish a formal AI governance policy covering sanctioned tools, permitted data classifications for AI processing, and agent permission standards within 30 days. Map AI asset discovery to your existing CMDB or asset management process. This gap exposes the absence of AI-specific controls in most existing GRC frameworks; use this as the trigger to add AI governance to your next compliance assessment cycle.

IR / Forensic Enrichment

Triage Priority	URGENT
Escalation Criteria	Escalate to a formal data breach investigation and engage legal counsel if the OAuth scope and file access log analysis (Step 2 evidence) reveals that any ungoverned AI agent had read or exfiltration-capable access to data stores containing PII, PHI, PCI-DSS-scoped data, or regulated intellectual property, particularly where data volume transferred to LLM API endpoints exceeds a de minimis threshold or where state breach notification laws (e.g., CCPA, NY SHIELD) or sector regulations (HIPAA, GLBA) apply.
Recovery Notes	Post-containment, maintain weekly telemetry-driven reconciliation of the AI asset inventory against the CMDB for a minimum of 90 days, as CrowdStrike field data indicates shadow AI re-emergence is rapid without enforced allow-listing at the proxy and endpoint management layer. Confirm that DLP inspection of LLM API prompt-layer traffic is functioning and classifying outputs — not just blocking domains — because users will pivot to sanctioned tools and attempt to submit the same sensitive data through approved channels. Retain all OAuth grant exports, DNS egress logs, and browser extension inventories from the remediation period for a minimum of one year to support any regulatory inquiry or downstream breach notification assessment triggered by the data classification review.

Forensic Artifacts	OAuth grant audit logs from M365 Entra ID and Google Workspace: export all grants with client_id, scope, consentType, and last-used timestamp — these identify which AI applications held persistent delegated access to corporate data under user identities with no security team visibility | DNS query logs and proxy egress logs filtered to LLM API hostnames (api.openai.com, api.anthropic.com, generativelanguage.googleapis.com, *.azure.com/openai): preserve per-user and per-device query frequency and data transfer volume to establish the scope of prompt-layer data submission and support T1567 (Exfiltration Over Web Service) triage | Browser extension manifest.json files from `%LOCALAPPDATA%\Google\Chrome\User Data\Default\Extensions\[ext_id]\` on Windows and `~/Library/Application Support/Google/Chrome/Default/Extensions/[ext_id]/` on macOS: preserve declared permissions (especially 'storage', 'tabs', 'clipboardRead', '') as evidence of the data access surface each unsanctioned AI extension had at the time of discovery | Sysmon Event ID 3 (NetworkConnect) logs correlating process name, user SID, and destination hostname for LLM API endpoints: these link specific ungoverned AI agent processes to specific user accounts, establishing the permission inheritance chain that enabled T1530 (Data from Cloud Storage) access | SaaS platform audit logs (M365 Unified Audit Log 'FileAccessed' and 'FileDownloaded' operations, Google Workspace Drive Audit 'download' events) correlated with the timestamps of outbound LLM API connections: this correlation is the primary evidence set for determining whether ungoverned agents staged and submitted sensitive documents to external LLM endpoints, and is required input for any breach notification threshold analysis

Per-Action IR Details

Step 1: Inventory — Run an AI asset discovery sweep across endpoints, SaaS integrations, browser extensions, and cloud-connected services. Do not rely on self-reported counts or IT procurement records. CrowdStrike's field data indicates those undercounts by 3x or more. Use endpoint telemetry and network egress logs to surface undeclared AI processes.

NIST Phase: Detection Analysis

Reference: NIST 800-61r3 §3.2 — Detection and Analysis: identifying the scope of affected systems and surfacing unknown assets before containment decisions can be made

Controls: NIST SI-4 (System Monitoring) — monitor for unauthorized AI processes and anomalous outbound connections to LLM API endpoints, NIST CM-8 (System Component Inventory) — extend inventory to include AI agents, copilots, and browser extensions as trackable system components, CIS 1.1 (Establish and Maintain Detailed Enterprise Asset Inventory) — discover all AI-capable assets including browser extension hosts and SaaS-connected endpoints, CIS 2.1 (Establish and Maintain a Software Inventory) — enumerate all installed AI browser extensions and agent software not currently in the authorized software list

Compensating: On Windows endpoints, run: `Get-Process | Where-Object {$_.MainWindowTitle -eq ''} | Select-Object Name, Id, Path` combined with `netstat -ano` filtered against known LLM API IP ranges for api.openai.com, api.anthropic.com, and generativelanguage.googleapis.com. Use osquery with query `SELECT name, identifier, path FROM browser_extensions WHERE name LIKE '%AI%' OR name LIKE '%GPT%' OR name LIKE '%Copilot%' OR name LIKE '%assistant%'` across enrolled endpoints. For SaaS OAuth exposure, export OAuth grant lists via Google Workspace Admin SDK or Microsoft Entra ID `Get-MgOAuth2PermissionGrant` PowerShell command and grep for AI vendor app IDs.

Evidence: Before initiating discovery, preserve: (1) a snapshot of current browser extension inventory from endpoint management (Intune, JAMF, or osquery browser_extensions table) to establish a pre-remediation baseline; (2) current OAuth grant exports from all SaaS platforms (M365, Google Workspace, Salesforce) showing granted scopes, grant dates, and authorized user counts; (3) 30-day DNS query logs filtered for *.openai.com, *.anthropic.com, *.googleapis.com/generativelanguage, *.azure.com/openai, *.cohere.ai to identify AI endpoints already in use; (4) network proxy or firewall egress logs showing data volume transferred to LLM API endpoints per user or device.

Step 2: Detection — Query endpoint logs for processes or browser extensions making outbound connections to known LLM API endpoints (api.openai.com, api.anthropic.com, generativelanguage.googleapis.com, and equivalents). Flag any agent operating under a standard user token accessing cloud storage (T1530) or submitting data to external web services (T1567). Review SaaS OAuth grant lists for AI applications with broad permission scopes.

NIST Phase: Detection Analysis

Reference: NIST 800-61r3 §3.2 — Detection and Analysis: correlating endpoint telemetry with network egress data to identify adversarial or ungoverned autonomous behavior patterns

Controls: NIST AU-6 (Audit Record Review, Analysis, and Reporting) — review endpoint and proxy logs for AI agent outbound connections to LLM APIs on a defined frequency, NIST AU-2 (Event Logging) — ensure logging is configured to capture browser extension network activity and process-level outbound connection events, NIST SI-4 (System Monitoring) — detect unauthorized data submission to external LLM APIs (T1567) and access to cloud storage under inherited user tokens (T1530), CIS 8.2 (Collect Audit Logs) — verify audit logging is enabled and collecting browser process network events, SaaS authentication logs, and OAuth grant activity

Compensating: Deploy Sysmon with a configuration that includes NetworkConnect events (Event ID 3) filtered on destination hostnames matching openai.com, anthropic.com, googleapis.com, azure.com. Query with: `Get-WinEvent -LogName 'Microsoft-Windows-Sysmon/Operational' | Where-Object {$_.Id -eq 3 -and $_.Message -match 'openai|anthropic|googleapis'}`. For OAuth scope analysis without a SIEM, use a Python script against the Microsoft Graph API or Google Workspace Admin SDK to flag any AI application granted Files.ReadWrite.All, Mail.Read, or equivalent broad-scope permissions. Apply the public Sigma rule 'proc_creation_win_lolbin_data_exfiltration' as a starting template adapted to flag curl/Invoke-WebRequest calls to LLM API endpoints.

Evidence: Preserve before flagging or alerting: (1) Sysmon Event ID 3 (NetworkConnect) records showing process name, user SID, destination IP and hostname for all LLM API destinations — this establishes which user accounts are running which AI agents; (2) Windows Security Event Log Event ID 4663 (Object Access) for any files accessed immediately before outbound LLM API connections, confirming T1530 data staging behavior; (3) SaaS audit logs (M365 Unified Audit Log, Google Workspace Admin Audit) for OAuth token issuance events to AI application client IDs within the last 90 days; (4) browser extension manifest files (located at `%LOCALAPPDATA%\Google\Chrome\User Data\Default\Extensions\[ext_id]\manifest.json`) for extensions with declared permissions including 'storage', 'tabs', 'clipboardRead', or '' that indicate broad data access capability.

Step 3: Eradication — Revoke OAuth grants for unvetted AI applications. Apply least-privilege to any AI agent or copilot that inherited broad user-level permissions. Block egress to unapproved LLM API endpoints at the proxy or firewall layer. Remove unsanctioned AI browser extensions via endpoint management policy.

NIST Phase: Eradication

Reference: NIST 800-61r3 §3.4 — Eradication: removing unauthorized components from the environment and eliminating the conditions that allowed ungoverned AI agents to persist and operate

Controls: NIST IR-4 (Incident Handling) — execute eradication as a formal phase of the incident handling capability, documenting each revoked grant and blocked endpoint, NIST AC-6 (Least Privilege) — reduce AI agent and copilot permissions to the minimum required for sanctioned functions, replacing inherited broad user-level tokens, NIST AC-2 (Account Management) — revoke OAuth grants for unvetted AI applications as unauthorized account-equivalent access relationships, NIST CM-7 (Least Functionality) — block egress to unapproved LLM API endpoints to eliminate unauthorized external service connections, CIS 5.4 (Restrict Administrator Privileges to Dedicated Administrator Accounts) — prevent AI agents from operating under elevated or overly-permissive user tokens, CIS 2.3 (Address Unauthorized Software) — remove unsanctioned AI browser extensions via endpoint management policy enforcement

Compensating: Revoke OAuth grants via PowerShell: `Remove-MgOAuth2PermissionGrant -OAuth2PermissionGrantId ` for M365 or equivalent Google Workspace Admin SDK call. Block LLM API egress on Windows hosts without a proxy by adding entries to the Windows Hosts file (`C:\Windows\System32\drivers\etc\hosts`) pointing api.openai.com, api.anthropic.com, and generativelanguage.googleapis.com to 127.0.0.1 as an immediate tactical block, then follow with a permanent firewall rule via `netsh advfirewall firewall add rule name='Block LLM APIs' dir=out action=block remoteip=`. Remove Chrome extensions without MDM using a Group Policy Object setting `ExtensionInstallBlocklist` with the specific extension IDs identified in Step 2.

Evidence: Before revoking any grant or blocking any endpoint, preserve: (1) full export of all OAuth grants with scopes, grant timestamps, and last-used timestamps from M365 Entra ID (`Get-MgOAuth2PermissionGrant -All`) and Google Workspace — this is your pre-eradication state record required for post-incident review; (2) screenshots or JSON exports of browser extension permission manifests for all extensions being removed, preserving the declared permission scopes as evidence of the exposure; (3) firewall and proxy logs covering the 90 days prior to blocking, capturing total data volume transferred to each LLM API endpoint per user — this establishes whether data exfiltration volume warrants escalation to a data breach investigation; (4) user account to AI agent mapping showing which service accounts or user tokens each ungoverned agent inherited permissions from.

Step 4: Recovery — Validate that your AI application inventory now reflects actual deployment counts via telemetry, not self-reporting. Confirm DLP controls cover prompt-layer traffic to external LLM APIs. Monitor OAuth grant lists and browser extension inventories on a recurring schedule (weekly minimum during remediation phase).

NIST Phase: Recovery

Reference: NIST 800-61r3 §3.5 — Recovery: verifying that the environment has returned to a known-good state and confirming that controls are in place to prevent recurrence of ungoverned AI deployment

Controls: NIST IR-4 (Incident Handling) — verify recovery completeness by confirming telemetry-driven inventory aligns with post-eradication expected state, NIST SI-7 (Software, Firmware, and Information Integrity) — use integrity verification to confirm no unsanctioned AI extensions or agents have been reinstalled post-eradication, NIST AU-6 (Audit Record Review, Analysis, and Reporting) — establish recurring review of OAuth grant lists and browser extension inventories at weekly frequency during remediation phase, NIST SI-4 (System Monitoring) — confirm DLP monitoring covers HTTPS traffic to LLM API endpoints including prompt-layer data submission, CIS 7.1 (Establish and Maintain a Vulnerability Management Process) — integrate recurring AI asset discovery into the vulnerability management cadence, CIS 8.2 (Collect Audit Logs) — validate that logging gaps identified during the discovery phase have been closed and that AI-related egress events are now captured

Compensating: Run a post-remediation osquery scan using `SELECT name, identifier, path, permissions FROM browser_extensions` and diff the output against the pre-remediation baseline captured in Step 1 to confirm removal and detect any reinstallation. For DLP coverage validation without a commercial DLP tool, configure a Squid proxy with SSL bump enabled and write an ACL that logs and blocks POST requests to /v1/chat/completions, /v1/messages, and equivalent LLM inference endpoints — inspect request bodies for PII patterns using a regex-based ICAP server such as c-icap with ClamAV. Schedule a weekly cron job or scheduled task to re-run the OAuth grant export and diff against the post-eradication state, alerting on any new AI application grants.

Evidence: During recovery validation, collect: (1) a fresh osquery browser extension inventory and OAuth grant export taken immediately after eradication, timestamped, to serve as the new clean-state baseline; (2) proxy or DNS logs from the 7 days post-eradication confirming zero successful connections to blocked LLM API endpoints — any hits indicate an eradication gap or policy bypass; (3) DLP policy coverage confirmation showing that outbound HTTPS POST traffic to LLM API hostnames is inspected, not tunneled blind through TLS — capture a sample DLP log entry showing prompt-layer traffic classification as evidence the control is functioning; (4) a dated reconciliation report comparing telemetry-derived AI asset count against prior self-reported count, documenting the specific gap as a measurable recovery milestone.

Step 5: Post-Incident — Establish a formal AI governance policy covering sanctioned tools, permitted data classifications for AI processing, and agent permission standards. Map AI asset discovery to your existing CMDB or asset management process. This gap exposes the absence of AI-specific controls in most existing GRC frameworks; use this as the trigger to add AI governance to your next risk assessment cycle.

NIST Phase: Post Incident

Reference: NIST 800-61r3 §4 — Post-Incident Activity: conducting lessons-learned, updating policies, and improving detection and governance capabilities based on findings from this incident

Controls: NIST IR-4 (Incident Handling) — update the incident handling capability to include AI governance gaps as a recognized incident class with defined response procedures, NIST IR-8 (Incident Response Plan) — revise the IR plan to incorporate AI asset discovery as a standing preparation activity and add AI-specific escalation criteria, NIST RA-3

(Risk Assessment) — add ungoverned AI deployments as an explicit risk category in the next risk assessment cycle, including agent permission scope and data classification exposure, NIST CM-8 (System Component Inventory) — integrate AI agents, copilots, and browser extensions into the CMDB as managed asset classes with defined approval workflows, NIST AC-2 (Account Management) — define AI agent permission standards as part of account provisioning policy, requiring explicit least-privilege scoping for all AI service accounts, CIS 1.1 (Establish and Maintain Detailed Enterprise Asset Inventory) — extend CMDB scope to include AI tools as a tracked asset class with discovery cadence defined, CIS 7.2 (Establish and Maintain a Remediation Process) — add AI governance control gaps to the risk-based remediation strategy with defined SLAs for sanctioning or removing AI deployments

Compensating: For teams without a GRC platform, create a lightweight AI governance register as a version-controlled CSV or Git repository documenting: tool name, vendor, OAuth scopes granted, data classifications permitted, approval status, and assigned owner. Automate weekly diff reporting by scheduling the osquery extension query and OAuth grant export as cron jobs with output compared against the register, emailing a delta report to the security team. Use the NIST AI RMF (AI 100-1) Govern function as a free reference framework to structure the policy — it is directly complementary to NIST CSF 2.0 and requires no licensing.

Evidence: Preserve as post-incident documentation: (1) the full timeline of AI asset discovery findings — specifically the ratio of telemetry-discovered AI deployments versus prior self-reported count — as quantitative evidence of the governance gap for executive and board reporting; (2) the pre- and post-remediation OAuth grant exports and browser extension inventories as before/after evidence packages supporting the lessons-learned report; (3) a documented mapping of which ungoverned AI agents had access to which data classifications (PII, PHI, financial, IP) based on the OAuth scopes and file access logs collected in Steps 1 and 2 — this is required to determine whether a data breach notification assessment is needed; (4) the DLP coverage gap analysis from Step 4, retained as evidence supporting the business case for AI-specific policy and tooling investment.

Detection Guidance

Primary detection surface is network egress and endpoint process telemetry. Query proxy or firewall logs for outbound HTTPS connections to LLM API endpoints from non-IT-managed processes or user-context agents. In SIEM, correlate user identity tokens (T1078) with cloud storage access events (T1530) and web service data transfers (T1567) to identify agents acting autonomously under inherited permissions. For SaaS environments, pull OAuth application grant reports from Google Workspace, Microsoft 365, and Salesforce and flag any AI application with scopes covering email, files, or calendar read/write that was not approved through IT procurement. For endpoints, query installed browser extension inventories for AI copilots and coding assistants not on an approved list. Falcon AIDR (if deployed) provides AI-specific behavioral detection; the Shadow AI Visibility Service (April 2026 launch) is designed specifically for inventory gap identification.

Framework Mappings

MITRE-ATTACK

- **T1078** — Valid Accounts
- **T1190** — Exploit Public-Facing Application
- **T1567** — Exfiltration Over Web Service
- **T1526** — Cloud Service Discovery
- **T1059** — Command and Scripting Interpreter
- **T1213** — Data from Information Repositories
- **T1195** — Supply Chain Compromise
- **T1530** — Data from Cloud Storage

NIST-800-53R5

- **AC-2** — Account Management
- **AC-6** — Least Privilege
- **IA-2** — Identification and Authentication (Organizational Users)
- **IA-5** — Authenticator Management
- **CA-8** — Penetration Testing
- **RA-5** — Vulnerability Monitoring and Scanning
- **SC-7** — Boundary Protection
- **SI-2** — Flaw Remediation
- **SI-7** — Software, Firmware, and Information Integrity
- **CM-7** — Least Functionality
- **SI-3** — Malicious Code Protection
- **SI-4** — System Monitoring
- **SA-9** — External System Services
- **SR-2** — Supply Chain Risk Management Plan
- **SR-3** — Supply Chain Controls and Processes
- **AC-3** — Access Enforcement
- **SC-28** — Protection of Information at Rest

OWASP-TOP10-2021

- **A01:2021** — Broken Access Control

HIPAA-SECURITY

- **164.312(a)(1)** — Access Control
- **164.308(a)(6)(ii)** — Response and Reporting

NIST-CSF-2

- **RS.CO-03** — Recovery activities and progress communicated

ISO-27001-2022

- **A.5.23** — Information security for use of cloud services

MITRE ATT&CK Mapping

Technique ID	Technique Name	Tactic
T1078	Valid Accounts	Defense-Evasion
T1190	Exploit Public-Facing Application	Initial-Access
T1567	Exfiltration Over Web Service	Exfiltration
T1526	Cloud Service Discovery	Discovery

Technique ID	Technique Name	Tactic
T1059	Command and Scripting Interpreter	Execution
T1213	Data from Information Repositories	Collection
T1195	Supply Chain Compromise	Initial-Access
T1530	Data from Cloud Storage	Collection

Sources

Source	URL	Tier
Blog	https://www.crowdstrike.com/en-us/blog/crowdstrike-shadow-AI-visibi...	T3
	https://www.crowdstrike.com/en-us/blog/crowdstrike-secures-growing-...	T3
	https://www.crowdstrike.com/en-us/blog/crowdstrike-stops-genai-data...	T3
	https://siliconangle.com/2026/03/23/crowdstrike-targets-ai-security...	T3
CrowdStrike Shadow AI Visibility Service | Reduce AI Footprint Risk	https://www.crowdstrike.com/en-us/services/ai-security-services/sha...	T3

Generated 2026-04-30 14:09 UTC by TJS Security Command Center