INTELLIGENCE BRIEFING
Security Command Center
TLP:CLEAR
2026-06-08 13:49 UTC

Shadow AI Governance Gap: 70% of Enterprise AI Runs Without Security Oversight as Agentic Risks Escalate

GOVERNANCE | MEDIUM | CVSS 5.0

SCC Item ID	SCC-GOV-2026-0045
Type	Governance
Severity	MEDIUM
CVSS Base Score	5.0
Affected Products	Enterprise SaaS platforms, LLM-integrated applications, agentic AI workflows, embedded AI copilots (vendor-agnostic); CrowdStrike Falcon AIDR and Shadow AI Visibility Service cited as detection/response tooling
Discovery Source	Rss:T1 Threatintel

Executive Summary

Seventy percent of enterprise AI deployments operate outside security team visibility, creating ungoverned exposure across data, identity, and automated actions. AI agents and embedded copilots process regulated data, execute API calls, and operate under borrowed credentials with no audit trail, no DLP enforcement, and no access controls. Organizations running agentic workflows face compounding risk: a single compromised or misbehaving agent can exfiltrate sensitive data, trigger unauthorized downstream operations, and leave no forensic record.

Technical Analysis

Shadow AI represents a structural governance failure across three attack surfaces. First, ungoverned data exposure: AI tools ingest regulated and sensitive data without DLP controls or data residency enforcement, mapped to CWE-200 (Exposure of Sensitive Information) and T1530 (Data from Cloud Storage) and T1567 (Exfiltration Over Web Service). Second, ungoverned action execution: agentic workflows execute multi-step API calls, file operations, and external communications without approval gates, mapped to CWE-284 (Improper Access Control) and T1059 (Command and Scripting Interpreter). Third, ungoverned identity: agents operate under borrowed human credentials or undocumented service accounts, bypassing least-privilege enforcement, mapped to CWE-778 (Insufficient Logging), CWE-20 (Improper Input Validation for prompt injection), and T1078 (Valid Accounts). T1199 (Trusted Relationship) applies where AI features are embedded in sanctioned SaaS platforms, making detection harder. T1048 and T1190 round out the technique set for exfiltration and prompt injection against AI-enabled applications. No CVE is assigned; this is a systemic governance condition. No

vendor patch exists. CrowdStrike has released a Shadow AI Visibility Service as a detection layer. This assessment draws from T3 sources (vendor blog, press release, trade press); the 70% statistic originates from Lenovo press release and is corroborated directionally by CrowdStrike research. Treat prevalence claims as indicative rather than audited. Authoritative governance baselines should reference NIST AI Risk Management Framework and CISA AI Security guidance (not yet cited in sources; recommended addition).

Action Checklist

1. Step 1: Discover - Audit all SaaS platforms for embedded AI features enabled by default (e.g., Microsoft 365 Copilot, Salesforce Einstein, Google Workspace Gemini). Disable AI features on platforms where no formal approval or data classification review has occurred. Query proxy and firewall logs for outbound API calls to known LLM endpoints (api.openai.com, api.anthropic.com, generativelanguage.googleapis.com, api.cohere.ai, huggingface.co). Reference: NIST AC-1 (AI Governance Policy) and CIS 4.4 (Firewall Configuration).
2. Step 2: Inventory - Create a formal AI asset registry cataloging all AI-integrated platforms, agentic workflows, and service accounts used for AI operations. Document data flows, data classifications, and access controls for each asset. Reference: CIS 1.1 (Asset Inventory) and NIST AC-2 (Account Management).
3. Step 3: Enforce - Implement approval gates for new AI tool adoption and API integrations. Require all AI-connected service accounts to be registered and scoped with least-privilege permissions per NIST AC-6. Enforce MFA on all accounts that can authorize AI workflows per CIS 6.3 and CIS 6.5. Enable logging on all AI-integrated platforms per NIST AU-2.
4. Step 4: Monitor - Query identity logs for service accounts or human accounts making high-frequency API calls to AI service endpoints. Review DLP telemetry for bulk data transfers to external AI APIs. Audit OAuth grants and third-party app permissions in SaaS admin consoles for AI-connected applications. Reference: NIST AU-6 (Audit Record Review, Analysis, and Reporting).
5. Step 5: Govern - Develop a formal AI governance policy establishing ownership, approval authority, data classification rules, and audit requirements for AI tool adoption. Establish a quarterly review cycle for shadow AI detection and policy updates. Document all control implementations against NIST AC-1 (Policy and Procedures) framework.

IR / Forensic Enrichment

Triage Priority	URGENT
Escalation Criteria	Escalate immediately to legal and compliance if evidence review confirms that ungoverned AI agents processed or transmitted records containing PII, PHI, or financial data subject to GDPR, HIPAA, or PCI-DSS — bulk data transfers to external LLM APIs meeting breach notification volume thresholds trigger mandatory reporting obligations independent of whether exfiltration was intentional.

Recovery Notes	Post-containment, maintain heightened monitoring of proxy egress logs and OAuth grant activity for a minimum of 30 days, as shadow AI re-adoption by end users following a visibility gap is common — users who relied on ungoverned tools frequently seek alternate paths once known endpoints are blocked. Verify that newly registered AI agent identities in the account inventory are being reviewed on the agreed cadence before the monitoring window closes. Confirm that DLP policies have been updated to classify AI API endpoints as data egress points, not merely external web destinations, so bulk submission events generate alerts rather than pass silently.
Forensic Artifacts	Microsoft 365 Unified Audit Log — CopilotInteraction workload events: captures user identity, timestamp, file or email objects submitted to Copilot, and response actions; this is the primary evidence source for ungoverned M365 Copilot usage and is not enabled by default in all tenants. | Azure AD / Entra ID OAuth2 grant logs — service principal credential issuance and token redemption events for app IDs matching known AI vendors (OpenAI, Anthropic, Google); timestamps on grant creation versus formal approval records establish the unauthorized access window. | Proxy or firewall egress logs — HTTP POST request records to api.openai.com, api.anthropic.com, and generativelanguage.googleapis.com with request/response payload sizes; large outbound payloads (>10KB per request) from service account source IPs indicate data being submitted to external LLMs rather than simple API health checks. | Salesforce Setup Audit Trail and Einstein Activity History — records AI feature enablement events, connected app OAuth approvals, and data object interactions by Einstein; these logs identify whether regulated Salesforce objects (Contacts, Opportunities with PII) were within scope of AI feature access. | Google Workspace Token Audit and Admin Activity logs — OAuth token grants to third-party AI applications with Drive, Gmail, or Docs scopes; cross-reference grant timestamps against the enterprise AI approval register to identify the full population of unauthorized integrations.

Per-Action IR Details

Step 1: Containment — Audit all SaaS platforms for embedded AI features enabled by default (e.g., Microsoft 365 Copilot, Salesforce Einstein, Google Workspace Gemini). Disable AI features on platforms where no formal approval or data classification review has occurred. Block outbound API calls to known LLM endpoints (api.openai.com, api.anthropic.com, generativelanguage.googleapis.com) via proxy or firewall for unapproved use cases. Reference: NIST AC-4 (Information Flow Enforcement) and CIS 4.4 (Implement and Manage a Firewall on Servers).

NIST Phase: Containment

Reference: NIST 800-61r3 §3.3 — Containment Strategy

Controls: NIST AC-4 (Information Flow Enforcement), NIST AC-20 (Use Of External Systems), CIS 4.4 (Implement and Manage a Firewall on Servers), CIS 4.5 (Implement and Manage a Firewall on End-User Devices)

Compensating: On Linux/macOS endpoints, add entries to /etc/hosts or configure a PAC file to null-route api.openai.com, api.anthropic.com, and generativelanguage.googleapis.com. On Windows, use Windows Firewall via PowerShell: `New-NetFirewallRule -DisplayName 'Block LLM APIs' -Direction Outbound -RemoteAddress (Resolve-DnsName api.openai.com).IPAddress -Action Block`. For SaaS audit without a CASB, export Microsoft 365 app consent grants via `Get-MgOauth2PermissionGrant` (Microsoft Graph PowerShell) and Salesforce Setup Audit Trail CSV; flag any OAuth grant issued to an AI app not on the approved list.

Evidence: Before disabling AI features, export Microsoft 365 Copilot interaction logs from the Microsoft Purview Audit portal (search for 'CopilotInteraction' workload events) and Google Workspace Admin Reports API for 'chat' and 'drive' events tied to Gemini. Capture current OAuth grant snapshots from each SaaS admin console — these are your baseline of what was authorized versus what was running. Document which platforms had AI enabled by default with no admin action, as this establishes scope of ungoverned exposure.

Step 2: Detection — Query identity logs for service accounts or human accounts making high-frequency API calls to AI service endpoints. Review DLP telemetry for bulk data transfers to external AI APIs. Audit OAuth grants and third-party app permissions in SaaS admin consoles for AI-connected applications. Enable logging on all AI-integrated platforms per NIST AU-2 (Event Logging) and CIS 8.2 (Collect Audit Logs). CrowdStrike Shadow AI Visibility Service provides dedicated detection primitives if Falcon is deployed.

NIST Phase: Detection Analysis

Reference: NIST 800-61r3 §3.2 — Detection and Analysis

Controls: NIST AU-2 (Event Logging), NIST AU-6 (Audit Record Review, Analysis, And Reporting), NIST AC-2 (Account Management), CIS 8.2 (Collect Audit Logs), CIS 5.1 (Establish and Maintain an Inventory of Accounts)

Compensating: Without SIEM, use PowerShell to parse Azure AD sign-in logs exported as JSON: filter for `appDisplayName` matching known AI services and `servicePrincipalName` patterns like '*openai*', '*anthropic*', '*gemini*'. For DNS-based detection on-prem, enable DNS debug logging on Windows DNS Server (Event ID 3008 in Microsoft-Windows-DNS-Client/Operational) and grep for ai service FQDNs using `Get-WinEvent`. Use Wireshark or tcpdump with a BPF filter (`host api.openai.com or host api.anthropic.com`) on an egress mirror port to capture payload sizes indicative of bulk data submission to LLM APIs.

Evidence: Collect Azure AD Unified Audit Log entries for OAuth2 token issuance to AI service principals (event type 'Add service principal credentials') before any revocation. Export Google Workspace Token Audit logs filtering on `api_name` for AI-related scopes (e.g., `gmail.readonly`, `drive.readonly` granted to third-party AI apps). Preserve DNS query logs showing frequency and volume of lookups to LLM provider domains — high-frequency repeated queries from a single service account identity are a key indicator of an agentic workflow running without human oversight.

Step 3: Eradication — Revoke undocumented service account credentials used by AI agents; issue new scoped credentials with least-privilege permissions per NIST AC-6 (Least Privilege) and CIS 5.4 (Restrict Administrator Privileges to Dedicated Administrator Accounts). Require all AI agent identities to be registered in your account inventory per CIS 5.1 (Establish and Maintain an Inventory of Accounts). Enforce MFA on all accounts that can authorize AI workflows per CIS 6.3 and CIS 6.5.

NIST Phase: Eradication

Reference: NIST 800-61r3 §3.4 — Eradication

Controls: NIST AC-6 (Least Privilege), NIST AC-2 (Account Management), NIST AC-3 (Access Enforcement), CIS 5.4 (Restrict Administrator Privileges to Dedicated Administrator Accounts), CIS 5.1 (Establish and Maintain an Inventory of Accounts), CIS 6.3 (Require MFA for Externally-Exposed Applications), CIS 6.5 (Require MFA for Administrative Access)

Compensating: In Azure AD, use `Get-MgServicePrincipal | Where-Object {$_.Tags -notcontains 'approved-ai'}` to identify unapproved AI service principals, then `Remove-MgServicePrincipalPassword` to revoke credentials. In Salesforce, navigate to Setup > Connected Apps OAuth Usage and click Revoke for any AI app not in the approved inventory. Issue replacement service account credentials with explicit scope restrictions — for OpenAI API integrations, replace organization-level API keys with project-scoped keys that carry data-residency and model-access restrictions. Document every new AI agent identity in a dedicated register (spreadsheet acceptable for small teams) with columns for owner, data access scope, approved use case, and review date.

Evidence: Before revoking credentials, capture the full permission scope of each service account or OAuth grant being revoked — specifically which Microsoft Graph API permissions (e.g., `Files.Read.All`, `Mail.Read`), Salesforce object permissions, or Google Workspace scopes were assigned. This establishes what data the AI agent could have accessed during its ungoverned operation period. Also snapshot the credential creation timestamps versus the date of any formal AI approval process to quantify the governance gap duration.

Step 4: Recovery — Validate that AI tool usage is now logged end-to-end with records meeting NIST AU-3 (Content of Audit Records) requirements: event type, timestamp, actor identity, data accessed, and action taken. Confirm DLP policies apply to AI-generated data transfers. Verify service account permissions have been scoped and documented. Monitor for re-emergence of shadow AI usage via proxy logs and OAuth audit trails per NIST AU-6 (Audit Record Review, Analysis, and Reporting).

NIST Phase: Recovery

Reference: NIST 800-61r3 §3.5 — Recovery

Controls: NIST AU-3 (Content Of Audit Records), NIST AU-6 (Audit Record Review, Analysis, And Reporting), NIST AU-9 (Protection Of Audit Information), NIST AU-11 (Audit Record Retention), CIS 8.2 (Collect Audit Logs)

Compensating: Configure a weekly automated PowerShell job that queries Microsoft 365 Unified Audit Log via `Search-UnifiedAuditLog -RecordType CopilotInteraction` and exports results to a protected log share — this creates the continuous AU-3-compliant record. For proxy-based shadow AI re-emergence detection, write a Sigma rule matching HTTP POST requests with `Content-Type: application/json` to AI provider IP ranges with request body sizes exceeding 50KB, which is indicative of bulk data submission rather than a single query. Forward proxy logs to a syslog collector (rsyslog is sufficient) with a 90-day retention policy to satisfy AU-11 minimums.

Evidence: Verify that newly enabled audit logs for Microsoft 365 Copilot, Salesforce Einstein Activity History, and Google Workspace Gemini are producing records containing all five AU-3 required fields before declaring recovery complete. Test by executing a known benign AI interaction and confirming the resulting log entry captures actor identity (not just a service account name — the human user who authorized it), the specific data object accessed (file name or record ID), and the AI action taken (query, summarization, generation). Absence of any field is a logging gap that must be remediated before monitoring is considered reliable.

Step 5: Post-Incident — This condition exposes three control gaps: absence of AI asset inventory (no mapped control in the provided knowledge base for AI-specific asset types — apply CIS 1.1 Establish and Maintain Detailed Enterprise Asset Inventory as the closest analog), absence of agentic workflow approval gates (NIST AC-5 Separation of Duties applies to automated action chains), and absence of audit trails for AI decisions (NIST AU-12 Audit Record Generation, AU-11 Audit Record Retention). Develop a formal AI governance policy under NIST AC-1 (Policy and Procedures) framework. Establish a recurring review cycle for new AI tool adoption.

NIST Phase: Post Incident

Reference: NIST 800-61r3 §4 — Post-Incident Activity

Controls: NIST AC-1 (Policy And Procedures), NIST AC-5 (Separation Of Duties), NIST AU-11 (Audit Record Retention), NIST AU-12 (Audit Record Generation), CIS 1.1 (Establish and Maintain Detailed Enterprise Asset Inventory), CIS 2.1 (Establish and Maintain a Software Inventory), CIS 7.1 (Establish and Maintain a Vulnerability Management Process)

Compensating: Build a lightweight AI asset register in a shared spreadsheet with mandatory fields: tool name, vendor, API endpoint, data classification of inputs/outputs, business owner, approval date, and next review date. Gate new AI tool onboarding using a two-person approval workflow (requestor plus security reviewer) enforced via a simple ticketing template — this operationalizes AC-5 Separation of Duties for agentic workflows without requiring a GRC platform. Schedule a quarterly OAuth grant review using the PowerShell and SaaS admin console exports established in Step 2, and retire any grant not re-approved by the owner within the review window.

Evidence: Conduct a lessons-learned review using the scope data collected in Steps 1–4: the list of platforms with AI enabled by default, the duration of ungoverned operation per service account, and the data classifications of content that AI features could have accessed. Quantify the blast radius — how many regulated data records (PII fields in Salesforce, email content in M365, files in Drive) were within scope of ungoverned AI access — as this number determines whether a breach notification assessment is required under GDPR, HIPAA, or state privacy law. This evidence package also forms the baseline metric for measuring improvement at the next review cycle.

Detection Guidance

Primary detection surfaces: (1) Proxy and firewall egress logs, filter outbound HTTPS to known LLM API hostnames (api.openai.com, api.anthropic.com, generativelanguage.googleapis.com, api.cohere.ai, huggingface.co) from endpoints and servers outside approved application inventories. (2) Identity and access logs, search for service accounts or human accounts with anomalous API call volumes to external AI services;

flag accounts with no MFA making AI API calls. (3) SaaS admin consoles, audit OAuth application grants for AI-connected third-party apps in Microsoft 365, Google Workspace, Salesforce, Slack, and GitHub; flag any app with broad data scopes (Mail.Read, Files.ReadWrite.All) connected to AI services. (4) DLP telemetry, alert on bulk transfers of files containing PII, financial data, or IP to AI API endpoints. (5) Endpoint telemetry, detect local model execution via CLI tools (ollama, llama.cpp, lm-studio processes). Behavioral indicators: large outbound payloads to AI endpoints during off-hours; service accounts initiating file reads followed immediately by external API calls; new OAuth grants appearing without change-management tickets. No specific IOC hashes or IPs are available for this governance-class condition.

Framework Mappings

MITRE-ATTACK

- **T1567** — Exfiltration Over Web Service
- **T1199** — Trusted Relationship
- **T1059** — Command and Scripting Interpreter
- **T1530** — Data from Cloud Storage
- **T1190** — Exploit Public-Facing Application
- **T1078** — Valid Accounts
- **T1048** — Exfiltration Over Alternative Protocol

NIST-800-53R5

- **CM-7** — Least Functionality
- **SI-3** — Malicious Code Protection
- **SI-4** — System Monitoring
- **SI-7** — Software, Firmware, and Information Integrity
- **CA-8** — Penetration Testing
- **RA-5** — Vulnerability Monitoring and Scanning
- **SC-7** — Boundary Protection
- **SI-2** — Flaw Remediation
- **AC-2** — Account Management
- **AC-6** — Least Privilege
- **IA-2** — Identification and Authentication (Organizational Users)
- **IA-5** — Authenticator Management
- **AC-3** — Access Enforcement
- **SI-10** — Information Input Validation
- **SC-28** — Protection of Information at Rest

OWASP-TOP10-2021

- **A01:2021** — Broken Access Control
- **A03:2021** — Injection

CIS-V8

- **6.1** — Establish an Access Granting Process
- **6.2** — Establish an Access Revoking Process
- **16.10** — Apply Secure Design Principles in Application Architectures
- **6.3** — Require MFA for Externally-Exposed Applications
- **8.2** — Collect Audit Logs

SOC2-TSC

- **CC6.1** — The entity implements logical access security software, infrastructure, and architectures over protected information assets

HIPAA-SECURITY

- **164.312(a)(1)** — Access Control
- **164.312(d)** — Person or Entity Authentication

ISO-27001-2022

- **A.8.26** — Application security requirements

NIST-CSF-2

- **DE.CM-01** — Networks and network services are monitored

MITRE ATT&CK Mapping

Technique ID	Technique Name	Tactic
T1567	Exfiltration Over Web Service	Exfiltration
T1199	Trusted Relationship	Initial-Access
T1059	Command and Scripting Interpreter	Execution
T1530	Data from Cloud Storage	Collection
T1190	Exploit Public-Facing Application	Initial-Access
T1078	Valid Accounts	Defense-Evasion
T1048	Exfiltration Over Alternative Protocol	Exfiltration

Sources

Source	URL	Tier
Blog	https://www.crowdstrike.com/en-us/blog/shadow-ai-hidden-risk-expand...	T3
	https://news.lenovo.com/pressroom/press-releases/70-enterprise-ai-u...	T3
	https://petri.com/shadow-ai-enterprise-threat-2030/	T3

Source	URL	Tier
	https://www.cio.com/article/4083473/shadow-ai-the-hidden-agents-bey...	T3
Introducing the CrowdStrike Shadow AI Visibility Service	https://www.crowdstrike.com/en-us/blog/crowdstrike-shadow-AI-visibi...	T3

Generated 2026-06-08 13:49 UTC by TJS Security Command Center