INTELLIGENCE BRIEFING
Security Command Center
TLP:CLEAR
2026-06-23 06:49 UTC

AI Agents and Policy Gaps: Palo Alto Networks Flags Identity and Access Control Risks in AI-Driven Cybersecurity

GOVERNANCE | MEDIUM

SCC Item ID	SCC-GOV-2026-0072
Type	Governance
Severity	MEDIUM
Affected Products	Enterprise environments deploying AI agents and agentic AI systems; broad applicability across sectors
Published	2026-06-21
Discovery Source	Gemini

Executive Summary

Palo Alto Networks has flagged a strategic governance gap: AI agents deployed across enterprise environments are not being held to the same identity and access management standards as human users, leaving organizations exposed to unauthorized privilege escalation, uncontrolled lateral movement, and incomplete audit trails. This is not tied to a specific CVE or active campaign; it is a structural risk emerging faster than policy and legislative cycles can respond. Any organization deploying agentic AI systems without formal IAM controls for those agents carries unquantified but material risk to its access governance posture.

Technical Analysis

Agentic AI systems operate as autonomous software entities that authenticate to services, execute actions, and chain tasks across systems, often with credentials or session tokens provisioned at deployment and rarely rotated. Unlike human accounts, AI agents frequently lack discrete identity records in IAM systems, bypass MFA requirements, and generate action logs that do not attribute causality to specific agent decisions. The attack surface includes: privilege escalation via over-provisioned agent service accounts, lateral movement through agent-to-agent or agent-to-API trust relationships, and audit trail gaps that obscure AI-initiated activity from SIEM correlation. No CVE, CWE, or active MITRE technique mapping is associated with this advisory. This is a governance and architecture risk signal, not an active exploitation disclosure.

Action Checklist

1. Inventory: Enumerate all AI agents deployed across the enterprise, including third-party SaaS tools with agentic capabilities, and register each as a discrete identity in your IAM system, apply CIS 1.1 (Establish and Maintain Detailed Enterprise Asset Inventory) and CIS 5.1 (Establish and Maintain an Inventory of Accounts) to AI agent identities explicitly.
2. Access Control: Audit service accounts and API tokens provisioned for AI agents; apply least-privilege scoping per CIS 3.3 (Configure Data Access Control Lists) and NIST AC-2 (Account Management) to verify agents cannot modify systems beyond their defined function, also map to D3-UAP (User Account Permissions).
3. Credential Hygiene: Rotate all long-lived credentials and API keys associated with AI agents on a defined schedule; enforce token expiry and re-authentication requirements, apply NIST IA-4 (Identifier Management) and NIST IA-5 (Authentication and Authorization); reference CIS 5.2 (Use Unique Passwords) and CIS 5.3 (Disable Dormant Accounts) for inactive agent identities.
4. Audit Logging: Verify that all AI agent actions generate attributable audit records meeting NIST AU-3 (Content of Audit Records) standards, confirm logs capture what action occurred, when, from which agent identity, and on which resource; ingest these logs into your SIEM and establish baseline behavioral profiles per NIST AU-6 (Audit Record Review, Analysis, and Reporting) and NIST SI-4 (System Monitoring).
5. Policy and Governance: Draft or update your AI governance policy to explicitly define IAM requirements for agentic systems, including identity registration, privilege scope, audit obligations, and incident response procedures for AI-initiated actions, reference NIST IR-8 (Incident Response Plan) to ensure AI agent scenarios are included in your IR playbooks, and align with NIST IR-4 (Incident Handling) for agent-specific escalation paths.

IR / Forensic Enrichment

Triage Priority	STANDARD
Escalation Criteria	Escalate to urgent if discovery of any AI agent service account reveals permissions exceeding its documented functional scope, active API key usage outside defined operational hours, or absence of any attributable audit records for an agent with write or delete access to sensitive data stores — any of these conditions indicates the structural risk has already materialized into a potential unauthorized access event requiring active incident response rather than governance remediation.
Recovery Notes	Because this advisory describes a governance gap rather than an active compromise, recovery is measured by closure of the structural control deficiencies: verify that 100% of AI agent identities are registered in IAM, all credentials have defined expiry and rotation schedules, and all agent actions produce attributable audit records before considering this risk closed. Monitor AI agent API call volumes and permission usage continuously for 90 days post-remediation to detect any behavioral drift indicating a previously unregistered agent remains active or a scoped-down agent is attempting actions outside its new permission boundary. Conduct a quarterly review cycle aligned with NIST IR-8 plan update requirements to reassess agent inventory as new agentic SaaS tools are adopted.

Forensic Artifacts	Cloud provider API call logs (AWS CloudTrail, Azure Unified Audit Log, GCP Cloud Audit Logs) filtered by service principal or assumed-role ARNs associated with AI agent identities — specifically events with write, delete, or IAM-modification actions that fall outside documented agent operational schedules | IAM credential reports showing API key creation dates, last-used timestamps, and permission scope for all non-human identities — gaps between documented agent deployment dates and key creation dates indicate unregistered agent provisioning | OAuth grant and service principal assignment records from the identity provider (e.g., Entra ID Enterprise Applications, Google Workspace OAuth grants) capturing which third-party SaaS tools with agentic capabilities were granted delegated permissions and what scopes were authorized | SIEM or log aggregation gaps for AI agent identity sources — date ranges where no agent telemetry exists despite the agent being operationally active represent blind spots where unauthorized lateral movement or privilege escalation would be undetectable | Agent runtime configuration files and environment variable stores (e.g., `.env` files, AWS Secrets Manager access logs, HashiCorp Vault audit logs) showing which credentials the agent was provisioned with at runtime and whether any credential access occurred outside the agent's normal execution pattern

Per-Action IR Details

Inventory: Enumerate all AI agents deployed across the enterprise, including third-party SaaS tools with agentic capabilities, and register each as a discrete identity in your IAM system — apply CIS 1.1 (Establish and Maintain Detailed Enterprise Asset Inventory) and CIS 5.1 (Establish and Maintain an Inventory of Accounts) to AI agent identities explicitly.

NIST Phase: Preparation

Reference: NIST 800-61r3 §2 — Preparation: Establishing IR capability requires knowing which systems and identities exist before an incident occurs; AI agents represent a non-human identity class absent from most asset and account inventories, making enumeration a foundational preparedness gap.

Controls: CIS 1.1 (Establish and Maintain Detailed Enterprise Asset Inventory), CIS 5.1 (Establish and Maintain an Inventory of Accounts)

Compensating: Run `az ad sp list --all` (Azure) or `aws iam list-roles` / `aws iam list-users` (AWS) to enumerate service principals and programmatic identities; cross-reference against a manually maintained spreadsheet tagging each identity as human, service account, or AI agent. For SaaS tools, audit OAuth grant lists in your identity provider (e.g., Google Workspace Admin Console > Security > API Controls) to surface agentic integrations that were provisioned without formal IAM registration.

Evidence: Before making any IAM changes: export a snapshot of all existing service principal assignments, API key metadata (creation date, last-used timestamp, scope), and OAuth grant records. In Azure, use `az ad app list --all > app_snapshot.json`; in AWS, `aws iam generate-credential-report` produces a CSV with key age and last-use data. This baseline is your forensic reference point for proving which AI agent identities existed prior to remediation and whether any were created outside change-control processes.

Access Control: Audit service accounts and API tokens provisioned for AI agents; apply least-privilege scoping per CIS 3.3 (Configure Data Access Control Lists) and NIST SI-7 (Software, Firmware, and Information Integrity) to verify agents cannot modify systems beyond their defined function — also map to D3-UAP (User Account Permissions).

NIST Phase: Preparation

Reference: NIST 800-61r3 §2 — Preparation: Reducing the blast radius of an AI agent identity compromise or privilege escalation event requires pre-incident access scoping; overprivileged AI agent service accounts are the primary structural risk identified in this advisory.

Controls: CIS 3.3 (Configure Data Access Control Lists), NIST SI-7 (Software, Firmware, and Information Integrity)

Compensating: For each AI agent service account, run `aws iam simulate-principal-policy` or review Azure role assignments with `az role assignment list --assignee ` to enumerate effective permissions. Compare against the agent's documented functional scope. Use the principle of least privilege to strip any permission not required for the agent's declared task — document removals in a change log. For on-premises environments, use PowerShell `Get-ADUser -Filter {ServicePrincipalName -ne '$null'} | Get-ADObject -Properties memberof` to surface service accounts with group memberships that may grant unintended lateral movement paths.

Evidence: Before modifying any access control assignment: capture a full export of current role bindings and effective permissions for each AI agent identity — `az role assignment list --all > role_assignments_pre_remediation.json` or AWS IAM Access Analyzer findings export. If an agent's permission scope is already broader than its documented function, treat this as a potential indicator of unauthorized privilege escalation and preserve the state as forensic evidence before scoping down. Document the delta between documented scope and observed permissions as a finding.

Credential Hygiene: Rotate all long-lived credentials and API keys associated with AI agents on a defined schedule; enforce token expiry and re-authentication requirements — apply D3-CRO (Credential Rotation) and D3-CH (Credential Hardening); reference CIS 5.2 (Use Unique Passwords) and CIS 5.3 (Disable Dormant Accounts) for inactive agent identities.

NIST Phase: Preparation

Reference: NIST 800-61r3 §2 — Preparation: Long-lived, non-expiring API keys assigned to AI agents are a structural vulnerability this advisory specifically identifies; establishing rotation schedules and expiry enforcement before an incident limits the window of exposure if an agent credential is exfiltrated or misused.

Controls: CIS 5.2 (Use Unique Passwords), CIS 5.3 (Disable Dormant Accounts)

Compensating: Use `aws iam list-access-keys` combined with `aws iam get-access-key-last-used` to identify AI agent API keys older than 90 days or with no recent usage — disable dormant keys immediately with `aws iam update-access-key --status Inactive`. For Azure, `az ad app credential list --id ` shows credential expiry dates; set `--end-date` on any credential lacking expiry. Script a monthly cron job or scheduled task to run these checks and output a CSV flagging keys exceeding your defined rotation threshold.

Evidence: Before rotating any credential: record the current key ID, creation timestamp, last-used timestamp, and associated service (from the IAM credential report or provider console) — this establishes forensic continuity proving which key was active during any prior period under investigation. If a key shows last-used activity inconsistent with the agent's documented operational schedule (e.g., activity at 03:00 UTC when the agent is supposed to be idle), treat this as a potential unauthorized use indicator and escalate to detection_analysis before rotating, to avoid destroying evidence of active misuse.

Audit Logging: Verify that all AI agent actions generate attributable audit records meeting NIST AU-3 (Content of Audit Records) standards — confirm logs capture what action occurred, when, from which agent identity, and on which resource; ingest these logs into your SIEM and establish baseline behavioral profiles per NIST AU-6 (Audit Record Review, Analysis, and Reporting) and NIST SI-4 (System Monitoring).

NIST Phase: Detection Analysis

Reference: NIST 800-61r3 §3.2 — Detection and Analysis: The incomplete audit trail problem flagged in this advisory means AI agent actions may be unattributable after an incident; establishing attributable logging for agent identities is the detection prerequisite for identifying unauthorized lateral movement or privilege escalation initiated by an AI agent.

Controls: NIST AU-3 (Content Of Audit Records), NIST AU-6 (Audit Record Review, Analysis, And Reporting), NIST SI-4 (System Monitoring)

Compensating: Enable AWS CloudTrail with `--include-management-events` and filter on `userIdentity.type = AssumedRole` or `userIdentity.arn` matching your AI agent service principal ARNs to isolate agent-sourced API calls. In Azure, query Unified Audit Log or Azure Monitor with `| where InitiatedBy contains "`. For teams without SIEM, export logs to a local file and parse with `jq` to extract agent identity, action, target resource, and timestamp — baseline normal operational hours and API call volumes manually in a spreadsheet. Deploy Sysmon on hosts where AI agents run locally, using a configuration that logs process creation (Event ID 1), network connections (Event ID 3), and file creation (Event ID 11) filtered to the agent process.

Evidence: This step does not alter live system state and requires no volatile pre-capture. However, before establishing baselines, archive a raw export of current logs covering the period since each AI agent was first deployed — this retrospective log pull is your forensic record for identifying whether unauthorized activity already occurred before audit controls were formalized. Specifically preserve: cloud provider API call logs filtered by agent service principal, any existing SIEM ingestion gaps (date ranges with no agent telemetry), and identity provider sign-in logs showing agent authentication patterns.

Policy and Governance: Draft or update your AI governance policy to explicitly define IAM requirements for agentic systems, including identity registration, privilege scope, audit obligations, and incident response procedures for AI-initiated actions — reference NIST IR-8 (Incident Response Plan) to ensure AI agent scenarios are included in your IR playbooks, and align with NIST IR-4 (Incident Handling) for agent-specific escalation paths.

NIST Phase: Post Incident

Reference: NIST 800-61r3 §4 — Post-Incident Activity: This advisory describes a structural policy gap, not an active exploit; the governance step is a lessons-learned and capability-improvement action that closes the policy vacuum enabling AI agent identity risks to persist — aligning with the post-incident function of improving organizational IR capability based on identified gaps.

Controls: NIST IR-8 (Incident Response Plan), NIST IR-4 (Incident Handling)

Compensating: A 2-person team can draft a minimum viable AI agent governance addendum using NIST IR-8's required plan components as a skeleton: add a section defining 'AI agent identity' as an account class, specify that all agents must be registered in the IAM inventory within 5 business days of deployment, and add a playbook entry for 'AI agent unauthorized action' that mirrors your existing service account compromise procedure. Store the policy in version control (Git) to maintain an auditable change history. Reference Palo Alto Networks' AI Security Reference Architecture publicly available guidance as a supplementary framework source during drafting.

Evidence: No volatile evidence capture is required for a policy drafting step. However, the policy itself should mandate that future AI agent incident investigations preserve: the agent's full API call history for the 30 days preceding detection, the agent's defined permission scope at time of incident versus observed effective permissions, any configuration changes made to the agent's IAM bindings in the 72 hours preceding the incident, and the agent's runtime logs showing what actions it took, on which resources, and in what sequence — establishing these evidence requirements in policy now ensures they are operationally available when an agent-initiated incident occurs.

Detection Guidance

No IOCs, exploit signatures, or specific detection rules apply to this advisory. Detection focus is behavioral and architectural. Query your IAM and SIEM for: (1) service accounts or API tokens with no associated human owner record and no MFA requirement, these are likely unmanaged agent identities; (2) accounts with broad privilege scopes that have never been used by a human login session; (3) API call sequences that chain across multiple systems without a corresponding human session context, which may indicate autonomous agent traversal. Apply NIST AU-12 (Audit Record Generation) to verify agent-initiated events are being captured, and NIST AU-6 (Audit Record Review, Analysis, and Reporting) to establish review cadence. Use NIST AU-6 (Audit Record Review, Analysis, and Reporting) to flag agent-associated accounts that exhibit unusual access patterns. The absence of agent entries in your IAM inventory is itself a detection signal, if you cannot enumerate your deployed agents, your logging coverage is incomplete.

Framework Mappings

NIST-800-53R5

- **AC-6** — Least Privilege

CIS-V8

- **5.4** — Restrict Administrator Privileges to Dedicated Administrator Accounts

SOC2-TSC

- **CC6.3** — Authorizes, modifies, or removes access

Sources

Source	URL	Tier
Palo Alto Networks says AI is reshaping cybersecurity policy and ...	https://x.com/EpicPlain/status/2068837269235028105	T3
How to Build a Generative AI Security Policy - Palo Alto Networks	https://www.paloaltonetworks.com/cyberpedia/ai-security-policy	T3
How AI Is Reshaping Cybersecurity and How Palo Alto Networks Is ...	https://www.cloudsyntrix.com/blogs/how-ai-is-reshaping-cybersecurit...	T3
What Is Generative AI in Cybersecurity? - Palo Alto Networks	https://www.paloaltonetworks.com/cyberpedia/generative-ai-in-cybers...	T3
Elusive threats, elastic defense: Securing AI at scale - IBM	https://www.ibm.com/thought-leadership/institute-business-value/en-...	T3

Generated 2026-06-23 06:49 UTC by TJS Security Command Center