Astrix

Securing the Agentic Frontier: Implementing CIS Controls with Astrix Security

Executive Summary

Astrix is an AI Agent Security Platform built to govern AI agents and the Non-Human Identities that power them. AI agents reason and act through NHIs: the API keys, service accounts, OAuth tokens, and MCP server credentials that define what an AI agent can read, write, and execute. Securing that access layer is how agent risk becomes governable. Astrix applies deterministic oversight to non-deterministic systems across the full lifecycle: discover, secure, and deploy.

In May 2026, the Center for Internet Security published three new Companion Guides extending the 18 CIS Critical Security Controls to AI environments. Astrix is the principal author on the LLM and AI Agent guides and a contributor on the MCP guide.

This paper maps Astrix's product capabilities to the CIS Safeguards called out in the AI Agent Companion Guide, tiered into three coverage levels: Primary, Supporting, and Not Relevant.

Coverage Map

Control	Coverage	Astrix role
1. Inventory of Enterprise Assets	Primary	Discovers and inventories AI agents, MCP servers, connected platforms
2. Inventory of Software Assets	Primary	Inventories the AI agent stack, flags deprecated and unofficial components
3. Data Protection	Supporting	Identity Graph for data flows, Secret Scanning for NHI credential exposure
4. Secure Configuration	Primary	Evaluates AI agent configurations against security baselines
5. Account Management	Primary	Full inventory and governance of NHIs and service accounts
6. Access Control Management	Primary	Business justification, RBAC evaluation, offboarding-driven revocation

Control	Coverage	Astrix role
7. Continuous Vulnerability Management	Supporting	Supplier threat intelligence, remediation workflows
8. Audit Log Management	Supporting	Anomaly detection on AI agent and NHI behavior
9. Email and Web Browser Protections	Supporting	Chrome extension inventory and risk analysis
10. Malware Defenses	Not relevant	-
11. Data Recovery	Not relevant	-
12. Network Infrastructure Management	Supporting	Identity Graph as architecture visualization, segmentation visibility
13. Network Monitoring and Defense	Supporting	Threat Cases, outbound webhooks to SIEM
14. Security Awareness and Skills Training	Not relevant	-
15. Service Provider Management	Primary	Supplier inventory, classification, continuous assessment, decommissioning
16. Application Software Security	Supporting	Third-party component mapping, severity rating, secure design checks
17. Incident Response Management	Supporting	Threat Cases, Action Center, post-incident forensic data
18. Penetration Testing	Not relevant	-

Primary Coverage

Visibility and Inventory (Controls 1 & 2)

Effective security of AI agents starts with knowing which ones exist, which NHIs they use, and which systems they reach.

Control 1: Inventory and Control of Enterprise Assets. Astrix automatically discovers and registers AI agents across platforms including Claude, ChatGPT, Copilot, AWS Bedrock, Azure AI Foundry, Google Vertex AI, n8n, and others (Safeguard 1.1). It captures critical metadata: agent name, version, connected data sources, risk level, and assigned human owner. Shadow AI and unapproved local agent tools are identified through AI Fingerprinting, with an Approval Status workflow to explicitly mark AI agents as Unapproved and generate findings like Forbidden Integration Installed (Safeguard 1.2). Active discovery uses existing EDR telemetry from Defender, SentinelOne, CrowdStrike, and others, so orchestration services and MCP servers exposed on the network are surfaced without direct code instrumentation (Safeguard 1.3).

Control 2: Inventory and Control of Software Assets. Astrix inventories the full AI agent stack: orchestration frameworks, MCP clients and servers, model SDKs, and associated suppliers (Safeguard 2.1). It tracks supplier origin and classification, flagging Deprecated Application, Deprecated Integration Type, and Legacy Application findings (Safeguard 2.2). Unauthorized components surface through Unofficial Application, Unreviewed Supplier, and Untrusted Supplier findings (Safeguard 2.3). Discovery is automated via API and EDR (Safeguard 2.4), and Governance Workflows enforce allowlisting by classifying components as Approved, Pending Review, or Unapproved (Safeguard 2.5).

Identity and Access (Controls 5 & 6)

AI agents act through non-human identities. These NHIs determine what an AI agent can read, write, and execute. Governing them is how agentic risk becomes governable.

Control 5: Account Management. Astrix maintains a complete inventory of service accounts, NHIs, API keys, and user accounts behaving as automated service accounts (Safeguards 5.1 and 5.5). Every identity asset in the inventory is assigned a human owner, and risk exposure is calculated for each. Shared or monolithic access is surfaced through Reused Credentials, Token Name Duplication, and Monolithic Credentials findings (Safeguard 5.2). High-privilege identities are flagged via Admin Privileges Granted and Super Admin Privileges Granted (Safeguard 5.4). A centralized governance pane tracks decentralized NHI credentials across platforms, mapping cross-platform access and identifying authentication methods (Safeguard 5.6).

Control 6: Access Control Management. Astrix explicitly inventories every third-party supplier, MCP server vendor, and AI platform connected to the environment via NHIs (Safeguard 15.1). Suppliers are classified and continuously assessed, with findings for Unreviewed Supplier, Untrusted Supplier, and reputation tiers from Highly Trusted to Compromised (Safeguards 15.2, 15.3, 15.5). The Threat Intelligence Database monitors providers for breaches and elevates the risk of associated integrations to Critical via the Compromised Supplier finding (Safeguard 15.6). Automated and manual workflows handle secure decommissioning of deprecated or risky integrations (Safeguard 15.7).

Configuration and Supplier Governance (Controls 4 & 15)

Control 4: Secure Configuration of Enterprise Assets and Software. Astrix evaluates AI agent configurations against security baselines, flagging Maker Mode in Copilot (where bots run using the creator's credentials) and insecure Username-Password Based Integrations (Safeguards 4.1 and 4.6). Through findings like Unused Consumer, Unused Credential, Tokens Unused for 6 Months, and Disabled Integration operational access may be continuously pruned (Safeguard 4.8).

Control 15: Service Provider Management. Astrix explicitly inventories every third-party supplier, MCP server vendor, and AI platform connected to the environment via NHIs (Safeguard 15.1). Suppliers are classified and continuously assessed, with findings for Unreviewed Supplier, Untrusted Supplier, and reputation tiers from Highly Trusted to Compromised (Safeguards 15.2, 15.3, 15.5). The Threat Intelligence Database monitors providers for breaches and elevates the risk of associated integrations to Critical via the Compromised Supplier finding (Safeguard 15.6). Automated and manual workflows handle secure decommissioning of deprecated or risky integrations (Safeguard 15.7).

Supporting Coverage: Unified Detection and Response

Astrix accelerates compliance by consolidating multiple CIS requirements into centralized operational workflows. Rather than managing controls in isolation, security teams leverage Astrix to automate detection and response across the following areas:

Threat Cases & Action Center (Safeguards 7.2, 7.7, 13.1, 16.2, 17.4, 17.5)

Astrix correlates anomalous agent behaviors, insecure configurations, and vulnerabilities into unified **Threat Cases**. These serve as the primary unit of investigation and are routed to the **Action Center**, which provides guided remediation, ticketing integration (Jira/ServiceNow), and chatbot workflows for rapid user communication or access revocation.

Behavioral Monitoring & Policy Thresholds (Safeguards 8.11, 12.2, 13.6, 13.11, 17.9)

The Anomaly Detection Engine establishes behavioral baselines for API and access logs to flag deviations such as "Anomalous Access" or "Sudden Activity." Organizations enforce custom security thresholds via **Source Access Policies**, which trigger alerts and block actions based on forbidden IPs, ISPs, or geographic regions.

Supply Chain & Vulnerability Intelligence (Safeguards 7.1, 7.4, 15.1-15.7, 16.5, 16.6)

Astrix integrates global threat intelligence to identify vulnerable or deprecated agent components. The platform flags "Unofficial" or "Suspicious" suppliers and assigns risk severity based on exposure and likelihood, ensuring the agent supply chain remains current and trusted.

Targeted Safeguard Coverage

- **Data Protection (Control 3):** Identifies overly broad ACLs (3.3), maps topological data flows via the **Identity Graph** (3.8), monitors transit encryption (3.10), detects cross-environment "contamination" (3.12), and provides identity-focused DLP via **Secret Scanning** for exposed NHI credentials (3.13).
- **Secure Design & Configuration (Controls 9, 12, 16):** Risk-scores browser extensions (9.4), exposes cross-platform lateral movement paths (12.4), and identifies design violations such as "Maker Mode" (Actions Defined Using Author Permissions) or broad organizational data sharing (16.10).
- **Forensics & Incident Management (Control 17):** Maintains historical activity timestamps and lifecycle tracking to reconstruct an agent's reasoning chains and "thought process" during post-incident reviews (17.8).

Out of Scope

Four CIS Controls fall outside Astrix's design focus and are acknowledged here for completeness.

Targeted Safeguard Coverage

- **Control 10: Malware Defenses.** Astrix tracks Compromised Suppliers and Malicious integrations based on threat intelligence but does not run an anti-malware execution engine, perform host-based file scanning, or manage OS-level anti-exploitation features.
- **Control 11: Data Recovery.** Astrix does not provide backup, retention, or data recovery capabilities.
- **Control 14: Security Awareness and Skills Training.** Astrix is a technical governance platform and does not provide human training programs.
- **Control 18: Penetration Testing.** Astrix is a continuous monitoring and posture management platform and does not conduct active penetration testing, red teaming, or exploit simulation.

Conclusion

The transition from static automation to autonomous AI agents shifts the security focus from code-level vulnerabilities to identity and behavioral governance. Aligning AI agent security with the CIS Controls lets enterprises extend a framework they already run into the layers that traditional assumptions didn't cover. Astrix provides the deterministic oversight and NHI governance that makes that alignment operationally real: AI agents accountable to human owners, operating within authorized boundaries, and maintaining a resilient posture across the enterprise.