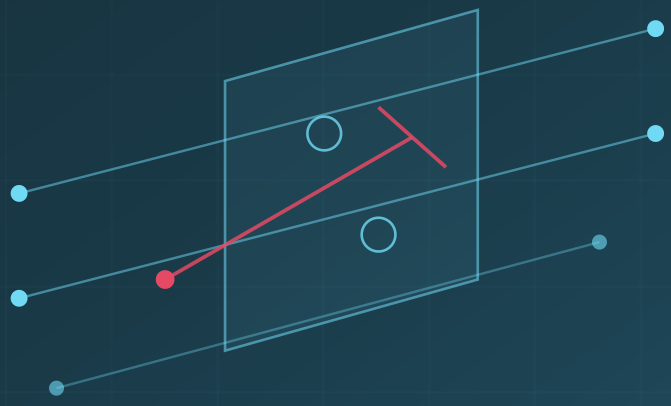
SecureAuth

AGENT AUTHORITY

Zero trust for AI agents

A guide for security and IAM leaders to what governs an agent's access at the moment it acts.



© 2026 SecureAuth Corporation.

Ebook

CONTENTS

Inside this ebook

01 **Agents act** 3
What changes when software decides and executes without a pause.

02 **The control gap** 4
Where an identity provider stops, and where an MCP gateway stops.

03 **Six control areas** 5
Discover, decide, detect, protect, meter and prove, and the capabilities that deliver them.

04 **Three enforcement points** 6
The device, the tool call and the model, under one policy set.

05 **MCP Authority** 7
Authorization depth, the catalog, self-service access, redaction and tool integrity.

06 **LLM Authority** 9
Prompts, data protection, model allowlists and budgets.

07 **Endpoint Authority** 10
Agents and MCP servers running on employee laptops.

08 **NHI Inventory** 11
The register of agents and non-human identities, each with an owner.

09 **Threats and findings** 12
Identity findings, behaviour anomalies and single-agent containment.

10 **Compliance and evidence** 13
Assessments, framework packs and audit-ready records.

11 **Deployment** 14
Integrations, standards, hosting and what changes in your agents.

12 **Getting started** 15
A first thirty days, and an evaluation checklist.

AUDIENCE

Security, IAM and technology risk leaders evaluating agent access controls.

COVERS

AI agents on employee devices, their tool calls over MCP and their model traffic.

SOURCES

The published Agent Authority product and capability pages. Early access is labelled.

© 2026 SecureAuth Corporation.

CHAPTER 1

Agents make autonomous decisions and act on them

An agent reads an instruction, decides what to do and executes, usually in under a second and usually against a credential issued for a human workflow.

Four things follow from that. An agent will execute a malicious instruction, because it treats instruction text as work to be done, including text that arrived inside a ticket, a document or a tool response. It improvises toward a goal, taking an unplanned route that can include operations nobody scoped on systems nobody intended. It acts as many identities, because runs share a deployment key and every run looks the same in the log. And it spends without approval, because token and tool consumption accrues with no metering and no human in the path.

All four have the same cause. Authorization happens once, at connection time, and nothing re-examines the action itself. The agent authenticates, receives broad reach, and every action after that inherits the same reach regardless of what it does, what data it carries or how the agent has been behaving.

A control model built for agents decides per action rather than per session, carries the delegating person through to the action, inspects responses as well as requests, and records the calls it refused.

KEY TERM

An *agent instance* is one run of an agent, bound to the person who authorized it, with its own identity rather than a shared deployment key.

WHY PER CALL

Consent happens once, at sign-in. An agent chooses a new action at each step, so authorization has to see each request and its arguments.

- **Compromised**
  Instruction text in a ticket, document or tool response is acted on as if it came from the operator.
- **Unaccountable**
  Runs share a key, so no action traces back to the person who delegated it.
- **Destructive**
  The agent optimises for an outcome and reaches systems the task never scoped.
- **Expensive**
  Spend accrues unmetered and arrives as one line that cannot be attributed.

THE CONTROL THAT FITS

Re-decide every call on the attributes it carries. The question is not whether this agent may reach this server. It is whether this agent, acting for this person, may call this tool with these arguments, given what the agent has done over the last ninety days.

© 2026 SecureAuth Corporation.

CHAPTER 2

02 Where an identity provider and an MCP gateway each stop

Most organisations start with one of two controls. Both are necessary. Neither covers agent actions on its own.

An identity provider authenticates the person

An IdP establishes who signed in. It does not govern what an agent does afterwards: which tool it calls, with what arguments, against which record, carrying what data back. Agent Authority sits downstream of an existing IdP and delegates authentication to it, adding per-run identity, per-call authorization, response filtering, spend attribution and a decision record.

An MCP gateway checks reachability

A typical gateway answers one question: is this server on the list and is this tool exposed. It then passes the call through and the response back, unchanged.

WORKS ALONGSIDE

SecureAuth, Okta, Ping Identity and Microsoft Entra ID. Groups come from the directory you already run.

ASSUMPTION

The enterprise identity provider stays authoritative for people and groups. Upstream systems keep enforcing their own entitlements.

	A TYPICAL MCP GATEWAY	AGENT AUTHORITY
GRANULARITY	Allow or deny per tool, not per operation, argument or field	**Re-decides every call on the attributes it carries**
IDENTITY	One key per deployment, so every run looks the same	**An identity per run, derived from the person and the process**
CONTAINMENT	Revoke the key and every agent stops	**One instance suspended and revoked in flight, every other agent left running**
RESPONSE	Forwards the response unchanged	**Filters the response before the agent reads it**
SPEND	Caps total spend	**Meters spend and attributes it to the people driving it**
EVIDENCE	Forwards the call and moves on	**Keeps the record, denied calls included**
AUTHENTICATION	Not its job	**Sits downstream of your IdP and delegates authentication to it**
COVERAGE	Sees only the agents pointed at it, blind to others	**Also enforced at the model and on the endpoint**

Table 1. The eight dimensions where a gateway and a control layer differ.

COVERAGE IS THE PRACTICAL LIMIT

A gateway only sees traffic routed through it. An agent installed on a laptop that calls a local MCP server, or a model endpoint directly, never appears in it. Enforcement has to exist at three points rather than one, which chapter 4 covers.

© 2026 SecureAuth Corporation.

CHAPTER 3

03 Control agent access and action, from discovery to record

Six control areas run in sequence. You cannot authorize what you have not discovered, and you cannot prove anything without a record. They operate as one system.

01 · DISCOVER

Agent identity and discovery

Inventory of MCP and LLM activity, discovery of shadow agents and non-human identities, and attribution of each agent back to a verified person.

02 · DECIDE

Runtime authorization and enforcement

Allow, deny or escalate per call. Destination control over where an agent may send data. Instant containment of a single instance.

03 · DETECT

Detection and risk

Behavioural anomaly detection against each agent's own baseline, identity risk findings, and assessment of what each agent can reach and change.

04 · PROTECT

Prompt and data protection

Context-aware redaction, sensitive data classification, prompt injection defence, and detection of secrets and credentials in flight.

05 · METER

Cost and operational controls

Rate limits and quotas, spend attribution and caps by team and person, and adoption and utilisation tracking.

06 · PROVE

Evidence and traceability

Agent Trail, the verdict record behind every decision, SIEM, SOAR and XDR integration, and visibility of enforcement traffic patterns.

How the areas are packaged

Six capabilities deliver these areas. Three are enforcement paths. Three are the register, the detection layer and the evidence layer behind them.

CAPABILITY	WHAT IT GOVERNS
MCP Authority	Tool calls made over the Model Context Protocol, scoped to the task
LLM Authority	Prompts and responses between agents and model providers
Endpoint Authority	Agents and MCP servers running on employee laptops
NHI Inventory	The register of agents, service accounts, keys and tokens, each with an owner
Agent Threats and Findings	Identity findings, behavioural anomalies, containment and SOC integration
Compliance and Risk Assessments	Framework mapping, risk ranking and audit-ready evidence

© 2026 SecureAuth Corporation.

CHAPTER 4

One authority for every path an agent takes

An agent reaches systems and data by three routes: the device it runs on, the tool calls it makes and the model it prompts. Agent Authority enforces at all three.

DEVICE PATH · ENDPOINT AUTHORITY

Actions that stay on the laptop

An agent a developer installed themselves, calling a local MCP server, never touching a corporate gateway. Discovery of AI apps, blocking of unsanctioned tools, and device posture signals that feed agent policy.

TOOL PATH · MCP AUTHORITY

Actions on a system of record

Creating a ticket, moving a deal, running a query. Tool authorization, task scoping and context-aware policy, with responses filtered before the agent reads them.

MODEL PATH · LLM AUTHORITY

Prompts leaving the company

Redaction of sensitive data, prompt injection defence, model allowlisting and token limits, applied before anything reaches a model provider.

One policy set across three paths

The same person, group membership and findings apply at all three routes. An agent whose identity has been flagged with a non-expiring secret can be kept off sensitive MCP servers. An agent running six times outside its normal call rate can have its next call denied wherever that call is heading. Policy is written once and enforced at whichever route the agent takes.

Each path reads the same person, instance and group context, and each writes to the same Agent Trail, so a review of one employee's agents covers every path at once.

ROLLOUT

Endpoint Authority ships as a sensor through Jamf or Intune. No browser change and no proxy change.

EVIDENCE

A blocked command on a laptop lands in the same record as a denied tool call at the gateway.

DEPLOYMENT IS CONFIGURATION, NOT CODE

There is no SDK and no agent rewrite. An MCP-capable agent is pointed at a brokered endpoint in its configuration file. A model-calling tool is pointed at a single endpoint with one line of config.

© 2026 SecureAuth Corporation.

CHAPTER 5

05 Authorize every MCP call, down to the arguments

MCP Authority is an MCP gateway and more: five pillars built as one system for enterprise scale, data privacy and fast adoption.

Authorization depth

Each agent run gets its own identity, and every tool call is authorized in real time. Responses are filtered before the agent sees them. The decision trace checks six things in order, then logs the rule that made the decision to Agent Trail.

DECISION TRACE

Person and group, agent identity for this run, server and tool, argument conditions, response data, live risk signals.

CHECKED ON EVERY CALL

Person and group → Agent identity for this run → Server and tool → Argument conditions → Response data → Live risk signals

A typical gateway checks server reachability and the tool list. Nothing below the tool name, nothing on the way back, and nothing about how the agent has been behaving.

The closed loop

Calls are recorded, anomalies and findings are detected, policy takes them into account, and enforcement applies on the agent's next call. An agent making six times its normal call rate can be denied. An agent identity holding a non-expiring secret can be kept off sensitive servers until that is fixed.

PILLAR 1

MCP Gateway

One gateway, or virtual gateways per business unit, each with its own endpoint and server set, managed from one console. Teams point an MCP-capable agent at the gateway URL in its config. Works with Claude Code, Cursor, Codex, GitHub Copilot, ChatGPT and Gemini.

PILLAR 2

MCP Catalog

More than 50 curated MCP servers, including vendors' official servers, grouped by category with their tools listed. Custom and internal servers are added by URL and inherit the same registry, access control, policy and audit.

PILLAR 3

Self-service access

Users browse tools by category and request the ones they need. Administrators work a Requests queue showing the tool, who asked and when. Users connect through MCP OAuth with their work login, so there are no shared API keys to distribute or rotate.

LOCAL SERVERS

MCP servers running on a user's own machine are handled by Endpoint Authority rather than the gateway, which is why the device is a separate enforcement point. Chapter 7 covers it.

© 2026 SecureAuth Corporation.

Data protection and tool integrity

Pillar 4 · Authorization and data protection

Policies read as sentences: agents, people and groups, MCP servers, tools, and optional argument conditions. Groups can come straight from Okta or Entra. One misbehaving agent can be cut off without affecting any other. Redaction policies filter server responses on the way back, using two mechanisms.

NO SHARED KEYS

Each user connects with their own work login over MCP OAuth, so access follows the directory rather than a distributed secret.

ARGUMENT CONDITIONS

Tool-level allow or deny forces a binary choice on tools that carry both routine and high-impact operations. Conditions permit the tool and constrain the operation.

DETERMINISTIC

Pattern filters

For known formats such as card numbers and dates of birth, where the shape of the data is predictable.

JUDGEMENT

AI judge

For described categories such as driver's licence or social security number. Administrators choose the behaviour if the judge fails.

Pillar 5 · MCP Security

CONTROL	HOW IT WORKS
Tool definition monitoring	If a tool's definition changes after approval, it is blocked until an administrator re-approves it. This covers rug-pull attacks and silent schema drift.
Authority Graph	Shows which groups reach which agents and servers, and which policies sit between them.
Agent Trail	Logs each call with the agent, the person, the server, the rule applied and redaction counts. Matched text is never stored, only counts.

DEPLOYMENT AND COMPLIANCE

Four hosting options

SecureAuth cloud (SaaS), dedicated or private cloud, on-premises, or fully air-gapped.

SOC 2 Type II and ISO 27001

Held for the service, separate from the compliance packs that map your own controls.

© 2026 SecureAuth Corporation.

CHAPTER 6

06 Access control for agents and AI models

LLM Authority sits between agents and model providers, checking each request against the agent, the person behind it and what it is sending.

Pillar 1 · Agent visibility

Each agent gets its own identity, tied to a verified person from your identity provider. Every prompt and response is logged with the agent, the person and the model reached. The audit trail is searchable and exportable to a SIEM. A live model-activity view shows token use by agent, and agents not on the approved list are flagged for review. One agent can be cut off without affecting the person who started it.

Pillar 2 · Data protection

Customer names, phone numbers and account details are detected in prompts and replaced with placeholders before they leave the company. A prompt can be redacted or blocked outright. Rules can differ by team, stricter for finance and support than for engineering. Content checks can run on a model endpoint you already approve.

DECISION TRACE

Identity, agent, model allowlist, redaction, budget, then allow or block. Checks run in real time on every request.

INTEGRATION

No agent rewrite and no SDK. One endpoint, one line of configuration per tool.

Pillar 3 · Models and budgets

Approved models are set by team or by person using your existing groups. Spend is shown by team, person and model, reconciled against actual provider costs. Requests stop when a budget is reached, and agents consuming tokens too quickly are cut off.

Pillar 4 · Prompt control

Prompts are classified by intent so rules can apply to each type. Prompt injection is detected, meaning instructions hidden in documents or tickets that try to steer the model. Alerts fire on unusual activity, such as a new model appearing in use or a sudden increase in spend.

SPECIFICATION	DETAIL
PROVIDERS	OpenAI, Anthropic, Google Gemini, Azure OpenAI, Amazon Bedrock
IDENTITY	From your identity provider. Logs export to your SIEM
CONTAINMENT	One agent cut off without affecting the person who started it

DECISION TRACE ON A MODEL REQUEST



Checks run in real time on every request. Integration is one endpoint and one line of configuration per tool, with no agent rewrite and no SDK.

© 2026 SecureAuth Corporation.

CHAPTER 7

07 Find and control AI agents on every laptop

Some agent actions never pass through a gateway. Endpoint Authority is a laptop sensor that finds the agents and MCP servers people run, ties each one to a verified person, and stops risky actions before they execute.

PILLAR 1

Shadow AI detection

Finds every agent on each laptop, including self-installed ones: Claude Code, Claude Desktop, Cursor, GitHub Copilot, Codex, Gemini CLI and desktop AI apps. Shows where each sends data and flags new destinations.

PILLAR 2

MCP server detection

Flags local MCP servers that are not on the approved list, and shows connections to MCP servers outside the company along with which agent opened each one.

PILLAR 3

Agent identity

Matches each device and agent to the person signed in through the identity provider. Rules use existing groups. When someone is offboarded, their agents lose access.

PILLAR 4

On-device control

Applies rules on the laptop before an action runs, blocking an unapproved MCP connection or a destructive git command while allowing an approved `gh pr create`.

One rule set across four agents

The same rules apply across Claude Code, Cursor, Copilot and Codex. Each decision is logged to Agent Trail with the person and the rule that produced it, so a blocked action on a laptop appears in the same record as a denied tool call at the gateway.

WHY THE DEVICE

Developer laptops are where agent adoption starts and where it is least visible. A gateway cannot see an agent that was never pointed at it.

SPECIFICATION	DETAIL
PLATFORM	macOS laptops, with a sensor status indicator in the console
ROLLOUT	Through existing device management: Jamf, Intune. No browser or proxy change
IDENTITY	Via your identity provider. Each laptop maps to a person once connected
DETECTION	Local MCP servers, remote MCP connections, each agent's data destinations

RULES APPLIED ON THE DEVICE, BEFORE AN ACTION RUNS

Block

A connection to an MCP server that is not on the approved list.

Block

A destructive git command issued by a coding agent.

Allow

An approved operation such as `gh pr create`, logged with the person and the rule.

AGENTS THE SENSOR RECOGNISES

Claude Code · Claude Desktop · Cursor · GitHub Copilot · Codex · Gemini CLI · Desktop AI apps

Offboarding a person removes access for the agents tied to them, because each device and agent maps to an identity in the directory.

© 2026 SecureAuth Corporation.

CHAPTER 8

08 Every agent and identity, each with an owner

NHI Inventory is a single register of AI agents and the service accounts, keys and tokens they use. It is the foundation the rest of Agent Authority builds on.

A policy based only on the user cannot tell two agents apart when they share the same credential. Identifying the acting agent, and the person accountable for it, comes before anything else. MCP Authority, LLM Authority, Endpoint Authority and Agent Threats and Findings all read from this register.

WHY FIRST

Policies based only on the user cannot distinguish two agents sharing one credential. The register makes the acting agent addressable.

READS FROM IT

MCP Authority, LLM Authority, Endpoint Authority and Agent Threats and Findings.

Pillar 1 · Agent inventory

Each agent is a named, versioned, registered non-human principal with an accountable owner, a stated purpose, a list of allowed tools and an owning team. The console shows totals for registered agents, agents active in the last 24 hours and tool executions in the last 24 hours, with runtime, version, tags, status and activity dates.

Pillar 2 · Discovery on devices EARLY ACCESS

Sensors on laptops, desktops and CI runners list devices with owner, last-seen time and active or revoked status. They report the AI agents running and flag processes missing from the catalog. Unregistered agents are added to the register. Each agent shows device and user counts, giving a measure of unmanaged AI use, and the AI domains it contacted.

Pillar 3 · Discovery in your directory EARLY ACCESS

Scheduled reviews of app registrations and agent identities in a connected directory such as Microsoft Entra. Findings are ranked by severity and name the rule that triggered, the identity and its kind, service principal or application. Each explains why it was flagged and what to do, and can be acknowledged or dismissed. Typical findings include tenant-wide admin consent, identities disabled by Microsoft, and secrets that never expire. A directory agents view lists agents with creation date, open-finding count, object type, app ID, owners and sponsors, which makes ownerless agents easy to spot.

Pillar 4 · Discovery sources

Connects to identity, endpoint and network providers you already run, including Microsoft Entra, Okta, CrowdStrike, Zscaler and F5, and flows everything discovered into one register. Each source shows its authentication status and when it was last checked.

WHAT THE REGISTER HOLDS

AI agents
Named, versioned and registered, with a purpose and allowed tools.

Service accounts
The non-human identities agents act through.

Keys and tokens
Including the ones that never expire.

Owners
An accountable person and an owning team for each entry.

© 2026 SecureAuth Corporation.

CHAPTER 9

09 Find your riskiest agents, and fix them first

Agent Threats and Findings surfaces risky agent identities and out-of-character activity, then lets you act on a single agent without disturbing the rest.

Identity findings

Scheduled reviews of app registrations, service principals and agent identities in your connected directory. Each finding carries a severity, the rule broken, the affected identity, its kind and when it was last seen. Findings sort by severity and filter by status, and each explains why it was flagged and what to do. Directory reviews are in early access.

Behaviour anomalies

Agent activity and spend are baselined over 90 days and ranked against peers. Each anomaly states what was observed against what is typical, how far it deviates, and shows an evidence chart of daily values against the expected range. It works across both MCP tool calls and LLM model use.

BASELINE

Ninety days of the agent's own activity and spend, ranked against peers rather than a fixed threshold.

CLOSED LOOP

Every call is recorded, anomalies are checked against the baseline, policy factors them in, and enforcement applies on the next call.

Containment, one agent at a time

ACTION	MECHANISM	EFFECT
Quarantine	Deny-all policy pushed to the enforcement sidecar	Outbound requests blocked within seconds
Revoke	Agent certificate revoked	mTLS connections rejected everywhere
Downscope	Specific permissions stripped	The agent keeps running with less reach

If an owner's account is disabled, you can assign a new owner or revoke the agent rather than leaving it orphaned.

Into the tools your SOC already watches

Agent audit events stream in real time via webhooks. Custom endpoints accept a Standard Webhooks envelope or a custom payload, each webhook reports delivery status as succeeded, retrying or failed, and setup is self-serve.

WEBHOOK DESTINATIONS

Slack · Microsoft Teams · PagerDuty · CrowdStrike Next-Gen SIEM · Datadog Log Management · Splunk HEC · Any HTTPS endpoint

Events carry the person, the agent, the rule and the result, so agent activity correlates with the rest of the estate in tools the SOC already runs.

IF THE OWNER LEAVES

An agent whose owner's account is disabled can be reassigned to a new owner or revoked. Either way it does not keep running unattributed, which is the state the register exists to prevent.

© 2026 SecureAuth Corporation.

CHAPTER 10

10 Audit-ready evidence for every AI agent

Which agents reach your most regulated services, how much they can change, and the record to prove what happened.

In-console assessments

Agent Access ranks identities by the regulated services their agents reach, with data class, tools used, sensitive call counts and severity. **Agent Risk** ranks them by how much their agents can change: whether they hold a complete exfiltration path, how many reached services are write-capable, and which tools can create, update, delete, send out or administer.

NHI Exposure Review

A one-off, read-only review run by a SecureAuth solutions engineer across identity providers, cloud accounts, code repositories, SaaS apps and AI agents. It inventories service accounts, API keys, tokens and their owners, ranks exposures by severity and returns a prioritised findings report.

AI Agent Risk Assessment

Uses telemetry the AI platforms already produce, pointed at SecureAuth. Setup takes about ten minutes, nothing is installed, and it does not sit in the agents' path.

AUDIT ENTRY

Agent ID, delegating user ID, tool name, parameters, timestamp and result, written at the time of each call.

REDACTIONS

Logged as match counts only. Matched text is never recorded.

Compliance packs

Map gateway safeguards to framework controls and show which are evidenced and which are not enforced. The coverage score counts only built-in evidence. Controls outside the gateway, such as security awareness training, are marked Organizational.

FRAMEWORK	CONTROLS	PACK VERSION
SOC 2	12	v1.0
HIPAA	12	v1.0
GDPR	8	v1.0
PCI DSS	12	v1.0
EU AI Act	7	v0.9

For the EU AI Act, most of what the gateway supplies is evidence (logs, oversight and monitoring) rather than a preventive control.

Audit-ready evidence

An evidence report per framework and period, with coverage, controls enforced, violations prevented, open items and a control matrix.

© 2026 SecureAuth Corporation.

CHAPTER 11

11 What it connects to, and where it runs

Agent Authority sits behind what you already own rather than replacing it. Authentication stays with your identity provider. Logs go to the SIEM your SOC already watches.

CATEGORY	NAMED INTEGRATIONS
IDENTITY	SecureAuth, Okta, Ping Identity, Microsoft Entra ID
AGENTS	Claude, ChatGPT, Gemini, Microsoft Copilot, Cursor, Replit
ENTERPRISE APPS	Salesforce, DocuSign, Slack, Microsoft 365, GitHub, Snowflake
SIEM / SOAR / XDR	Splunk, CrowdStrike, Datadog, Microsoft Sentinel
DISCOVERY	Microsoft Entra, Okta, CrowdStrike, Zscaler, F5
DEVICE MGMT	Jamf, Intune, for the Endpoint Authority sensor

Standards

OAuth 2.1, OpenID Connect, SPIFFE/SVID, mTLS, token exchange, and the Model Context Protocol.

What changes in your agents

Nothing. There is no code change and no SDK. An MCP-capable agent is pointed at the brokered endpoint in its configuration. A model-calling tool is pointed at a single endpoint with one line of config. Enforcement then covers tool calls, endpoint devices and LLM prompts.

HOSTING

SecureAuth cloud, dedicated or private cloud, on-premises, or fully air-gapped.

CERTIFICATION

SOC 2 Type II and ISO 27001, held for the service.

CREDENTIALS

No standing credentials for agents. Agents use OAuth and short-lived tokens rather than long-lived keys. Each run gets its own identity tied to the person and the process that started it, and an instance can be revoked or downscoped quickly. Over-privileged, stale and shared credentials are flagged.

© 2026 SecureAuth Corporation.

CHAPTER 12

12 Start this week

The sequence below follows the product's own order: see what exists, decide what it may do, then prove it.

	STEP	WHAT IT INVOLVES
01	**Establish the register**	Add a discovery source and let NHI Inventory populate. Look for agents nobody registered, identities with no owner, and secrets that never expire.
02	**Run a read-only risk analysis**	The free NHI Exposure Review covers identity providers, cloud accounts, repositories, SaaS apps and agents, and returns a prioritised findings report. Nothing sits in the agents' path.
03	**Pick the first enforcement route**	Developer laptops point to Endpoint Authority. Agents acting on systems of record point to MCP Authority. Data leaving for a model provider points to LLM Authority.
04	**Point one agent at the endpoint**	One line of configuration, no rewrite and no SDK. Review the decision trace in Agent Trail before writing any deny rules.
05	**Write policy as sentences**	Agents, people and groups from Okta or Entra, servers, tools and argument conditions. Begin with broad rules and narrow them using what Agent Trail records.
06	**Turn on redaction and the loop**	Pattern filters for known formats, the AI judge for described categories, anomaly baselining running behind both.
07	**Connect the SOC and pick a framework**	Stream audit events to Splunk, Sentinel, Datadog or CrowdStrike, then open the compliance pack for whichever framework your next audit uses.

EVALUATION CHECKLIST

Granularity
Can one operation on a tool be denied without denying the tool itself.

Containment
Can a single instance be contained while every other agent keeps running.

Evidence
Does the audit record include denied calls, not only successful ones.

Coverage
Is enforcement available on the device and at the model, not only at the gateway.

NEXT STEPS

Fourteen days in your own Agent Authority tenant, with a read-only NHI Exposure Review included.

FREE TRIAL
secureauth.com/free-trial/agent-authority

TALK TO AN ARCHITECT
secureauth.com/demo

DOCUMENTATION
docs.secureauth.com/ai

PRODUCT
secureauth.com/products/agent

© 2026 SecureAuth Corporation.