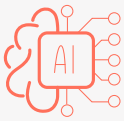
CATO
NETWORKS

CISO Six-Step Plan for Secure AI Transformation

The biggest productivity unlock in decades is also the main source of uncontrolled risk in the enterprise. AI's influence extends across every business function and is central to innovation and growth. As gains materialize, dependency deepens and adoption continues to increase.

But the pace of AI adoption has outrun security controls that were never built to govern models and their data flows. Every interaction with an AI tool, app, or agent moves, processes, and exposes data across enterprise boundaries that are largely invisible to traditional monitoring.

CISOs are managing competing pressures:



boards demanding AI-driven efficiency and competitive advantage

business units seeking speed and workflow compression

regulators requiring granular evidence of how AI is governed and audited

This eBook is a six-step plan to address those pressures.

It gives CISOs and their teams a path to establish visibility, control, and governance across AI in use, development, and production, without slowing the business.

Why securing AI is different

Traditional security models depend on predictable environments. Applications have defined boundaries, data flows through documented paths, and control points sit in known locations. AI violates every single one of those.

Approving a shortlist of sanctioned AI tools is necessary, but it is not sufficient to secure and govern AI. Risk sits in what flows into AI tools and models, what comes back out, and where outputs are allowed to influence decisions. A single interaction can combine user inputs and prompts, internal enterprise data, external model calls, and automated actions.

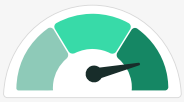
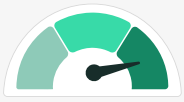
And prompts and outputs vary by request. Autonomous AI agents add another layer by acting across systems without a human gate check.

AI interactions can surface multiple risk types at once: data exposure, runtime misuse, inaccurate or unverified output, compliance gaps, and third-party dependencies. Treating these separately creates gaps, delays, and inconsistent accountability.

AI risk types CISOs must manage

As organizations operationalize AI, the same risk types tend to surface. These define the CISO's scope of AI risk responsibility.

The table below maps each risk to a strategic response and indicates the level of influence security teams typically have using existing authority, tools, and controls.

Risk Type	Strategic Response	Degree of Security Control
Data Privacy & Leakage Sensitive data shared with AI tools through prompts, uploads, and connected workflows.	**Inventory and Enforce Responsible AI Usage** Discover all AI activity and apply data-handling policies.	 **HIGH** Most organizations can extend adjacent controls, though effectiveness depends on integration and visibility.
Model Accuracy and Integrity Inaccurate or misleading output influencing business decisions.	**Establish AI Risk Governance** Introduce validation and decision governance for AI-generated output that influences business outcomes.	 **HIGH** Requires both technical controls and process alignment.
Runtime Security Threats Threat vectors like prompt injection, misuse, model poisoning, and malicious agents.	**Implement Guardrails** Apply inline guardrails. Enforce runtime controls and red-teaming of AI interactions.	 **HIGH** Addressable when runtime controls are consistently enforced across environments.
Compliance and Regulation Lack of auditability and policy enforcement.	**Stay Ahead of Requirements** Track AI usage, enforce acceptable use policies, block non-compliant AI interactions, and generate audit evidence aligned to regulatory expectations.	 **MEDIUM** Standards continue to evolve.
Third Party and AI Supply Chain Risk Opaque model behavior and data handling practices.	**Enable Responsible AI Development** Add AI-specific checks into existing SDLC and vendor review.	 **MEDIUM** Depends on vendor transparency and contractual controls.

Six steps to secure AI transformation

Risk categories define the scope. The next move is execution.

The six foundational steps below translate scope into action. They show how CISOs can operationalize AI security across usage, development, and production, applying consistent controls across users, locations, clouds, and AI interactions.

STEP 1

Gain visibility into AI usage

See it to govern it

AI use across the enterprise is often user-sanctioned and invisible to IT, introducing shadow AI risk. This includes employees sharing sensitive data with GenAI tools, developers experimenting with open-source models, and teams enabling AI features inside approved SaaS platforms without security oversight.

Shadow AI creates exposure across data protection, compliance, and third-party risk because these interactions frequently occur outside traditional monitoring and control points.

Visibility into AI usage is the starting point for control. This includes:

- Which AI applications and agents are in use, sanctioned and unsanctioned
- Who is accessing them, and from where
- What data flows into and out of AI systems
- Where interactions occur across cloud, branch, and endpoint

CISO starting point

Inventory every AI workflow across the organization: internal, external, approved, or shadow. Map what you find to your data classification policies.

STEP 2

Establish secure AI Posture Management

Fix risks before they reach production

Before models go live or AI agents touch production data, reduce avoidable exposure in configuration and access. AI workflows span data ingestion, model training, inference, and integration. Each stage can contain vulnerabilities: exposed APIs, over-permissive access, or unsafe configurations.

Effective AI Security Posture Management (AI-SPM) must continuously analyze:

- AI service configurations and access policies
- Misconfigured API calls or plugin permissions
- Sensitive data linked to models or agents
- Anomalous or unapproved connections to public AI systems

The goal is to proactively detect and remediate AI risk before deployment.

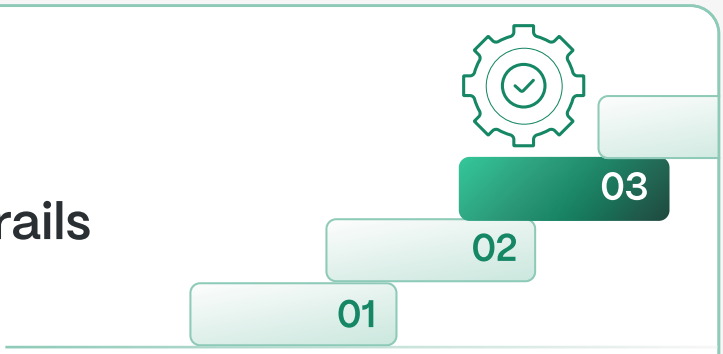
CISO starting point

Move AI security checks into the build and release pipeline. Integrate AI-SPM checkpoints into DevOps and MLOps so security validation becomes a release requirement rather than a post-launch audit.

STEP 3

Implement runtime protection and guardrails

Stop attacks in real time



Once AI is active, security shifts from static configuration to dynamic defense. New risks like prompt injection, data exfiltration, and model manipulation can only be addressed while AI is running.

Effective runtime protection requires continuous guardrails on AI interactions as they happen:

- Enforce policies for sensitive data and model access
- Detect and block prompt injection, jailbreak attempts, and unsafe or off-policy model responses
- Restrict anomalous behavior and unauthorized API calls triggered by autonomous agents

Where possible, controls should be applied inline so enforcement happens as data moves. API and integration-based approaches offer deeper application layer visibility, but inline enforcement is the most consistent way to govern sensitive data across users, applications, third-party, and homegrown AI systems in real time. The goal is protection across AI deployment models without introducing separate control silos.

CISO starting point

Define guardrail policies around sensitive data usage and high-risk AI interactions. Test and validate enforcement using simulated AI attacks/controlled test scenarios to ensure controls respond before sensitive data exits the environment.

STEP 4

Integrate AI risk into the SOC

Operationalize AI security

AI introduces a new category of security telemetry. If those signals live in a separate AI security console or outside existing SIEM/SOAR workflows, response slows and operational complexity increases.

Integrating AI risk into the SOC allows security teams to investigate incidents without introducing new tools or consoles.

This requires correlating AI activity with identity, data, and network context, including:

- Sensitive data shared with public AI systems
- AI-driven anomalies across users, regions, or tenants
- Repeated model misuse or policy violations

Alerts should feed directly into existing SIEM/SOAR pipelines like Splunk, Microsoft Sentinel, or CrowdStrike Falcon, enabling faster triage and response.

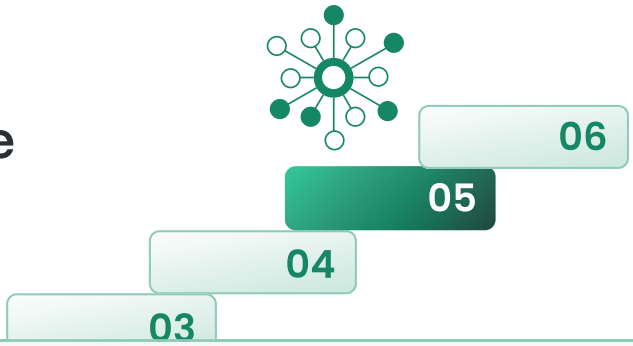
CISO starting point

Review your SOC playbooks. Add AI-related alert categories and response workflows to avoid treating AI misuse as noise in your event stream.

STEP 5

Strengthen governance and compliance

Make AI auditable



Regulatory requirements around AI continue to evolve. Organizations will be expected to demonstrate responsible use, accountability, and auditability across AI systems.

Effective governance requires:

- Monitoring AI usage against compliance and acceptable-use policies
- Capturing interaction metadata and audit evidence for every AI interaction
- Generating audit-ready reporting aligned with recognized frameworks like NIST AI Risk Management Framework (AI RMF), OWASP Top 10 for LLMs, and EU AI Act requirements for public facing applications

This turns governance into a continuous process rather than resource-heavy, periodic audits.

CISO starting point

Align your internal policies with recognized frameworks and create evidence pipelines now. Producing audit artifacts reactively after a regulatory inquiry is not a defensible position.

STEP 6

Govern how AI output gets used

Protect decision integrity

Even a technically secure AI system can still create business risk through incorrect or misleading output. When that output feeds into business workflows, inaccuracies become operational risk, not just a quality issue.

Decision integrity depends on identifying where AI output influences business, legal, financial, operational, or other critical decisions, and applying review or verification rules that match the consequence.

No security product can determine with certainty whether a specific AI output is factually correct or safe for business use. The control point is the process: knowing which decisions rely on AI output and putting appropriate review guardrails around them.

This requires CISOs to:

- Identify the decisions where AI output should never go unchecked by a human
- Build verification steps into high-impact or regulated workflows that consume AI-generated content
- Use existing policy enforcement and audit mechanisms to track where AI output is shared and consumed
- Measure whether review requirements are being followed

This extends AI security beyond protecting data, into governing how AI output is used inside the business.

CISO starting point

Identify the decisions where AI output carries real consequence. Require verification or human review for those outcomes and measure compliance.

The importance of architecture

AI security should not be treated as a standalone layer bolted onto existing security controls. Different types of AI risk appear across prompts, data shared with models, outputs returned, and the systems those outputs connect to.

Application controls can provide deeper context within specific AI tools, browser or endpoint controls help monitor user-driven activity, and API-based controls support better visibility into certain interactions and integrations. Each of these has value.

But AI risk does not stay inside one layer. A single AI workflow can move across users, SaaS applications, APIs, models, internal data sources, and automated actions. As those interactions cross environments, siloed controls create gaps in visibility and enforcement.

That is the architectural problem.

The most defensible control point is the one that can inspect data in motion across those environments and enforce policy before sensitive information leaves trusted boundaries. The network provides the strongest enforcement point because it sits in the path of traffic moving between users, applications, data sources, and third-party AI services. That makes it the place where AI security should live.

The Bottom Line

AI adoption will continue to accelerate. The security challenge will grow with it. CISOs need governance that performs under real usage.

At Cato Networks, we believe the same architectural principles that made SASE (Secure Access Service Edge) transformational for networking and security convergence, now define the journey to secure AI transformation: visibility, control, and governance delivered from a unified, cloud-native platform.

With Cato AI Security, natively built into Cato's unified SASE architecture, CISOs don't have to choose between innovation and control. Cato enables teams to govern AI usage, protect data in motion, and reduce risk across the enterprise without adding another disconnected control layer:

- See what AI tools and data flows exist across the organization
- Lock down configuration and posture before anything moves to production
- Put controls around decisions that depend on AI output
- Route AI signals into the SOC so analysts can act on them
- Generate audit evidence before the board or regulatory bodies ask
- Enforce runtime guardrails

Want to see Cato AI Security in action?

Cato experts can help map your AI exposure points, demonstrate Cato controls, and show you what reporting looks like in your environment.

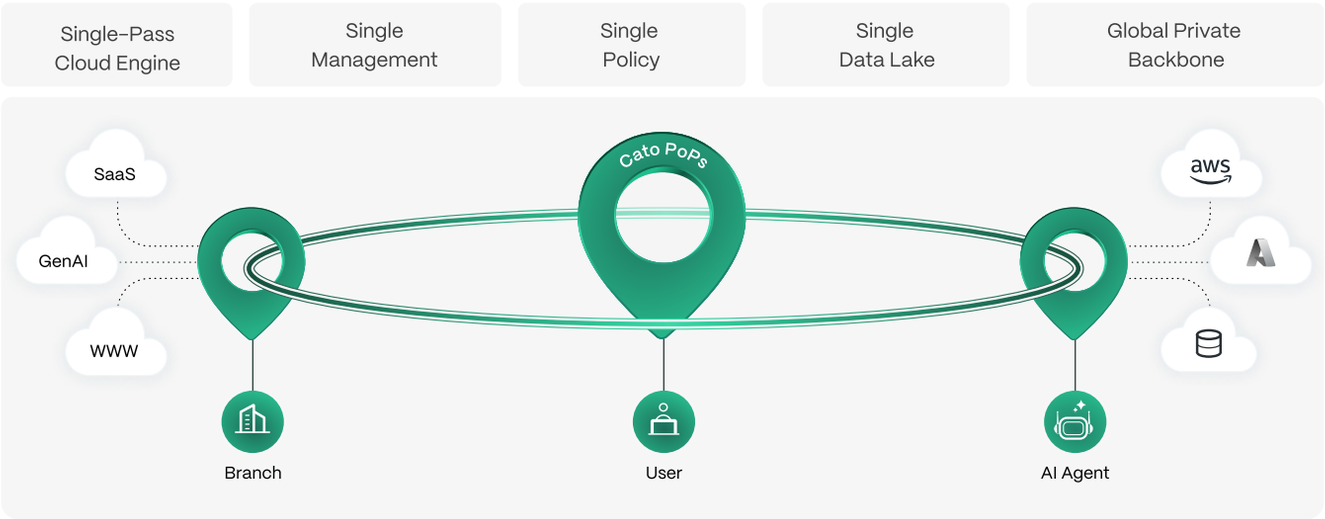
Schedule a personalized demo

About Cato Networks

Cato Networks, a leader in SASE and AI security, delivers secure, zero-trust access everywhere to thousands of customers worldwide. Built for organizations operating across all cloud and hybrid environments, the Cato Platform unifies networking, security, and access, providing them as elastic, modular capabilities that organizations can easily adopt and grow over time. Cato combines the Cato Cloud, a purpose-built global network, with simplified operational experience, all delivered across a robust, AI-driven platform. With Cato, organizations modernize confidently, operate with greater resilience, and innovate faster, without added complexity or risk.

The Cato SASE Platform

Start your networking and security transformation wherever the business need is most urgent. From there, tap into the full power of the Cato SASE Cloud: purpose-built and self-optimizing global cloud with a single-pass engine, single data lake, and single policy enforcement.



Cato SASE Platform

Solutions

Next Gen Networking

Universal ZTNA

Zero Trust Security

AI Security

SASE Convergence

Cato Use Cases

Network Transformation
MPLS to SD-WAN Migration
Global Access Optimization
Hybrid Cloud & Multi-Cloud

Business Transformation
Vendor Consolidation
Spend Optimization
M&A and Geo Expansion

Security Transformation
Secure Hybrid Work
Secure Direct Internet Access
Secure Application and Data Access
Incident Detection and Response

Secure AI Adoption
Public AI Usage
Private AI Applications

Contact Us