GovCIO MEDIA & RESEARCH

JULY 2026

DeepDives

BUILDING AI GUARDRAILS

INSIDE:

Table of Contents

AI Success Depends on Modern Infrastructure, Strong Governance

Federal leaders must modernize infrastructure, strengthen data governance and remove shadow AI to safely deploy advanced AI capabilities at scale.

The federal mandate for AI integration is an active operational requirement, but rushing to deploy advanced algorithms within a fragmented legacy IT infrastructure creates profound systemic risks, data leakage vectors and governance vulnerabilities.

Federal IT estates are characterized by historical data silos, outdated legacy hardware and a highly distributed workforce. Hastily adopting commercial AI utilities outside structural boundaries creates severe risk profiles, including accidental exposure of personally identifiable information (PII) and protected health information (PHI).

Observability starts with visibility. You can't fix what you don't see. Many organizations still have siloed teams using siloed tools with siloed data in siloed environments, which causes blind spots. To achieve full visibility, a federal agency must adopt a structured, five-phase framework to systematically modernize infrastructure, establish unyielding data governance, eradicate shadow AI and scale safely toward autonomous systems.

Phase 1: Foundation and Governance Establishment

The initial phase focuses entirely on engineering the administrative machinery and institutional oversight structures mandated by OMB memorandum M-25-21 and the NIST AI Risk Management Framework (AI RMF). Without an authoritative governance posture, subsequent technological integrations risk duplication, compliance failure and unchecked risk propagation.

Structural Leadership and Cross-Functional Oversight

Agencies must formalize executive leadership by explicitly designating a chief AI officer (CAIO). The CAIO acts as the central authority responsible for

Photo credit: Summit Art Creations/Shutterstock

spearheading mission innovation, optimizing infrastructure investments and identifying systemic risk parameters across the enterprise.

Concurrently, agencies must convene a formal AI governance board. This board must not sit isolated within IT; it requires cross-functional participation from executive stakeholders spanning IT, cybersecurity, data architecture and legal counsel to evaluate proposals through multi-disciplinary lenses.

Comprehensive Inventory Baseline and Structural Alignment

Before deploying new algorithmic models, an agency must maintain a comprehensive understanding of its operational baseline. This requires cataloging all current, legacy and planned AI use cases into a centralized, transparent inventory. This cataloging mechanism serves to illuminate critical intersection points where emerging AI components interface with legacy mission systems and core enterprise data assets. In parallel, internal policies regulating IT infrastructure, cybersecurity boundaries and data privacy must be aggressively revisited and rewritten to natively incorporate AI-specific controls, preventing policy fragmentation.

Implementing the NIST “Govern” Core

Operationalizing the foundational “govern” function of the NIST AI RMF demands a systematic shift in organizational culture. Agencies must formally codify acceptable use policies for employees, mandate explicit classifications for data sensitivity and institute an operational risk taxonomy. This taxonomy must cleanly separate routine, lower-risk administrative automation from “high-impact” applications that affect public welfare or regulatory determinations, ensuring appropriate tiers of scrutiny are applied downstream.

(ctd.)

Photo credit:
Tee11/Shutterstock

Phase 2:
IT Estate Modernization and Data Architecture

Algorithmic models perform only as well as the data architectures that feed them and the compute environments that host them. To transition away from restrictive data silos and brittle, outdated technical debt, agencies must architect resilient, cloud-native environments designed to support rapid AI prototyping while maintaining rigorous security postures.

Secure Cloud Compute Fabrics

AI model execution, fine-tuning and inference require modern, specialized compute fabrics. Workloads must be transitioned away from legacy environments into secure, FedRAMP-approved computing environments.

These environments must feature absolute logical segmentation, leveraging virtual networks and micro-segmentation to isolate AI models. Furthermore, role-based access control (RBAC) must be strictly enforced, ensuring that only authorized personnel and verified system processes can interact with model parameters and training sets.

Unified Data Architectures and FAIR Principles

Maximizing model accuracy requires tearing down structural data silos. Agencies must focus on consolidating structured, unstructured and semi-structured data repositories into optimized enterprise data lakes and data

Photo credit: Gorodenkoff/Shutterstock

warehouses. Consolidating data into unified enterprise data lakes without establishing strong metadata tagging increases the risk of model degradation and compliance challenges.

Implementing findable, accessible, interoperable, reusable (FAIR) principles ensures data assets remain structured, searchable and secure across the entire federal enterprise lifecycle. Data must be expertly curated, cleaned and appended with rich, descriptive metadata tags. This rigorous data preparation directly mitigates structural bias, optimizes model training velocity and ensures traceability.

Controlled Experimentation Sandboxes

To foster organic mission innovation without endangering the live production environment, agencies must engineer isolated experimentation sandboxes. These secure, pre-authorized environments grant internal developers access to vetted, open-source AI foundation models, specialized software tools and compliant libraries. Developers can safely test capabilities, run stress tests and prototype applications within a ring-fenced boundary before moving code into the official deployment pipeline.

Phase 3: Capability Building and Low-Risk Experimentation

With governance established and modernized cloud architecture in place, agencies can safely introduce AI capabilities to the active workforce. This phase prioritizes driving immediate administrative value through low-risk services while systematically upskilling personnel and identifying shadow IT risks.

Scaling Low-Risk Enterprise Services

Initial enterprise value is most efficiently realized by rolling out commercial off-the-shelf (COTS) generative AI assistants and standardized productivity solutions. Deploying these tools to knowledge workers streamlines tedious

Photo Credit: Digitala World/ Shutterstock

back-office workflows, accelerates document summarization and optimizes routine administrative tasks. All software deployments must progress through a formalized, phased development pipeline — advancing sequentially through sandbox, development, pilot and production milestones — with rigorous compliance check-ins at every phase transition.

Workforce Literacy and Strategic Upskilling

The successful integration of AI requires an AI-ready workforce. Agencies must execute tiered AI literacy programs designed to upskill employees based on their operational roles.

General staff should receive training centered on basic AI awareness, capability boundaries and strict adherence to acceptable use guidelines. Conversely, technical cadres — such as data scientists, software engineers and cloud architects — must receive advanced, highly specialized training regarding model optimization, safe fine-tuning practices and algorithmic risk mitigation.

Eliminating Shadow AI Threat Vectors

One of the most immediate vulnerabilities facing federal networks is shadow AI: the unauthorized use of public, consumer-grade AI platforms by employees seeking to optimize their daily tasks. To stop this threat vector, agencies must aggressively deploy forward proxies and implement network monitoring and data loss prevention controls. These security measures detect and block unauthorized data transfers to unvetted public models, ensuring that sensitive data assets, including PII and PHI, never leak outside the sanctioned security boundaries of the agency.

(ctd.)

Photo Credit:
DC Studio/Shutterstock

Phase 4: High-Impact Deployment and Rigorous Risk Mitigation

As agencies successfully mature past basic administrative tools, they will inevitably embed AI into mission-critical environments and “high-impact” use cases. These include systems where AI outputs serve as a primary or substantial basis for decisions directly affecting public safety, emergency response, civil rights or critical infrastructure. At this juncture, risk guardrails must become absolute and unyielding.

Pre-Deployment Testing and Impact Assessment

No high-impact AI system can be authorized for live operational environments without undergoing extensive pre-deployment validation. Agencies must mandate exhaustive, multi-stage testing regimens, including adversarial red-teaming designed to intentionally trick, exploit or bypass model logic. Furthermore, program managers must compile comprehensive AI impact assessments. These formal documents must explicitly quantify model output quality, weigh expected mission benefits against potential societal or operational harms and define concrete mitigation strategies for identified vulnerabilities.

Hybrid Privacy-Enhancing Technologies

Protecting data integrity at scale requires a defense-in-depth posture utilizing advanced privacy enhancing technologies (PETs). For real-time, operational inference models where latency is a critical factor, agencies should apply data obfuscation techniques and differential privacy to shield underlying source records. Conversely, for highly sensitive, high-security offline computations, intensive cryptographic frameworks such as fully homomorphic encryption must be reserved and deployed, allowing mathematical analysis to be executed directly on encrypted data states without ever exposing the raw information. (ctd.)

Photo Credit: puhhha/Shutterstock

Human-in-the-Loop Architecture and Reversibility

Automated systems must never supersede human accountability in high-impact scenarios. Agencies must implement human-in-the-loop safeguards that cannot be bypassed, along with mandatory human review checkpoints.

Crucially, system architects must engineer a deterministic reversibility capability directly into the software fabric. This programmatic capability ensures that any automated or AI-assisted mission decision can be immediately halted, overridden and entirely undone by human operators the moment an anomaly or unsafe output is detected.

Phase 5: Mature Enterprise Scaling and Autonomous Agentic AI

The zenith of the federal AI journey is the creation of a fully scalable, highly interconnected ecosystem capable of leveraging autonomous capabilities safely and efficiently across the entire scope of the enterprise.

The Centralized Federal AI Marketplace

To eliminate redundant development costs and accelerate cross-agency collaboration, organizations should establish a centralized AI marketplace. This

“Adopting AI across the federal government is an architectural challenge, not merely a software procurement task.”

internal repository acts as a secure clearinghouse where developers can list, showcase and securely share reusable machine learning features, custom code snippets, pre-trained weights, and verified, clean datasets. This approach drives rapid scaling not just within an agency, but across the broader landscape of the federal government.

Securing Agentic AI via Zero Trust Architecture

The next evolutionary phase of automation is agentic AI: advanced systems capable of executing multi-step, autonomous tasks and leveraging external software tools across disparate databases without manual human intervention. In an agentic paradigm, traditional human-centric identity and access management approaches become insufficient.

To secure autonomous agent boundaries, agencies must integrate agentic AI directly into their zero trust architectures. This requires assigning distinct, cryptographic machine identities to every autonomous agent. Furthermore, agencies must enforce strict metadata-based, least-privilege access controls. This guarantees that an autonomous agent can only read, write or execute commands within tightly defined parameters, neutralizing the risk of lateral privilege escalation or cascading autonomous system failures.

Automated Continuous Monitoring and Drift Detection

Post-deployment management requires operationalizing the “manage” and “measure” functions of the NIST AI RMF through automated, continuous monitoring infrastructure. Agencies must utilize advanced analytics platforms to track real-time model drift, identify data anomalies and rapidly flag hallucinations or confabulations before they skew mission decisions.

The monitoring ecosystem must feature automated feedback loops, ensuring that any detected operational incidents or anomalies are fed directly

Photo Credit: MMD Creative/ Shitterstoick

back into model retraining and fine-tuning cycles. This creates a self-healing, continuously optimizing enterprise infrastructure.

Conclusion:
A Call to Action for Federal Leadership

Adopting AI across the federal government is an architectural challenge, not merely a software procurement task. True mission success demands a rigorous adherence to structured modernization timelines, strict data governance and proactive security paradigms. By executing this five-phase roadmap, federal IT decision-makers can confidently fulfill the statutory requirements of OMB M-25-21, implement the safeguards of the NIST AI RMF, and deliver a secure, resilient and highly advanced AI-driven enterprise.

Photo Credit:
VideoFlow/Shutterstock

GovCIO MEDIA & RESEARCH

DeepDives

Inside NIST's Risk Management Framework

The National Institute of Standards and Technology's AI Risk Management Framework provides voluntary guidance on designing, developing, deploying and using AI systems. The framework is part of NIST's broader goal to build trust in AI systems.

MAP

Identifies the AI system's purpose, operational context, stakeholders and potential impacts to establish the scope of risk.

MEASURE

Analyzes and validates risks using metrics for trustworthiness through testing, red teaming and bias assessments.

AI RISK MANAGEMENT FRAMEWORK

GOVERN

Cultivates a risk-aware organizational culture, establishing policies, procedures and accountability to guide the other three functions.

MANAGE

Implements controls based on measurements, prioritizing risks based on impact and likelihood.

Source: NIST

GovCIO MEDIA & RESEARCH

PARTNER INTERVIEW

Symantec. by Broadcom | Carbon Black. by Broadcom | carahsoft.

Data Visibility Remains the Biggest AI Security Gap

Federal agencies can balance AI adoption with security by improving data visibility, governance and control over sensitive information.

Garrett Lee
Regional Vice President, Public Sector, Enterprise Security Group, Broadcom

How are you seeing federal agencies navigate the rapid rise of generative AI, and where do the biggest gaps lie when it comes to protecting sensitive data and maintaining control?

Lee Federal agencies are moving with urgency to adopt generative AI, driven by administration priorities and a recognition that they need to keep pace with technological change. However, the risk calculus for government differs from that of the private sector. While commercial organizations often take a "pilot first, govern later" approach, federal agencies don't have that luxury. They manage highly sensitive information — including classified data, controlled unclassified information (CUI), citizen data, law enforcement information and HIPAA-regulated health records — where mistakes can have national security implications and erode public trust.

That creates a tension between providing employees with access to productivity-enhancing tools and ensuring data

remains protected within established policy boundaries. Blanket bans are rarely effective and often drive the use of unauthorized tools, creating shadow AI risks.

The biggest gap we see today is data visibility. Many government leaders acknowledge they don't have a complete understanding of where all sensitive data resides, how it is classified or who has access to it. Because generative AI systems rely on broad access to data, those visibility gaps create significant risk. Agencies may unintentionally expose sensitive information to AI systems without fully understanding how that data could be used, retained or shared.

What defines an effective approach for agencies seeking to balance AI adoption with robust data protection, and which capabilities or strategies are proving most impactful?

Lee An effective approach treats security as an enabler rather than a barrier. Agencies need a data-centric security strategy that allows them to innovate while maintaining control over sensitive information.

Traditional network perimeters are no longer enough. Security controls must follow the data itself, with visibility across endpoints, networks and data environments. That starts with strong cyber hygiene, including data classification, tagging, encryption and data loss prevention (DLP) capabilities that continuously monitor how information is accessed and shared.

Agencies also need practical guardrails. Rather than slowing innovation, they should establish approved AI environments with clearly defined use cases, segmented access controls and continuous monitoring. There should be clear policies governing what information can be shared with public models versus

"An effective approach treats security as an enabler rather than a barrier."

— Garrett Lee, Regional Vice President for Public Sector at Broadcom's Enterprise Security Group

private or agency-approved AI systems.

It's equally important to focus on both AI inputs and outputs. Challenges such as hallucinations, poisoned training data, AI-enabled phishing campaigns and increasingly sophisticated malware are expanding the threat landscape. To address these risks, agencies should leverage behavioral analytics, endpoint protection and strong identity governance to continuously assess and validate their security posture.

This is where zero trust principles become critical. Capabilities such as DLP and extended detection and response (XDR) provide visibility into how data moves across the environment, helping security teams identify risks, enforce policy and respond quickly to emerging threats.

As agencies continue to evolve their AI strategies, what changes or developments are you most focused on, and how do you see the role of cybersecurity and data governance shifting over the next year?

Lee Over the next year, I expect agencies to move beyond early experimentation and take a more deliberate approach to AI architecture and governance. Organizations will increasingly focus on questions such as which models are appropriate for specific types of data, where inference can safely occur and how human oversight should be incorporated into mission-critical workflows.

I also expect growing interest in sovereign AI and private AI deployments, particularly among defense, intelligence and highly regulated civilian agencies. As organizations evaluate the costs, security requirements and data exposure risks associated with public cloud environments, many will look more closely at private cloud and self-hosted options.

As a result, cybersecurity teams will play a broader role in AI governance. Their responsibilities will extend beyond protecting infrastructure to overseeing data lineage, validating model integrity and managing access to sensitive information throughout the AI lifecycle.

Ultimately, AI is forcing agencies to address data management and governance challenges that have existed for years. The agencies that succeed won't necessarily be the ones that move fastest. They'll be the organizations that build trust, visibility and control alongside innovation.

Photo Credit: Gorodenkoff/Shutterstock

Cyber Ready for the AI Era.

Empowering security teams with the visibility, control, and speed needed to outpace today's most sophisticated threats.

As threats evolve with AI, so does your defense. Symantec and Carbon Black equip teams to detect faster, respond smarter, and recover stronger. In a world where threats never sleep, Symantec and Carbon Black keep your organization one step ahead.

Learn More

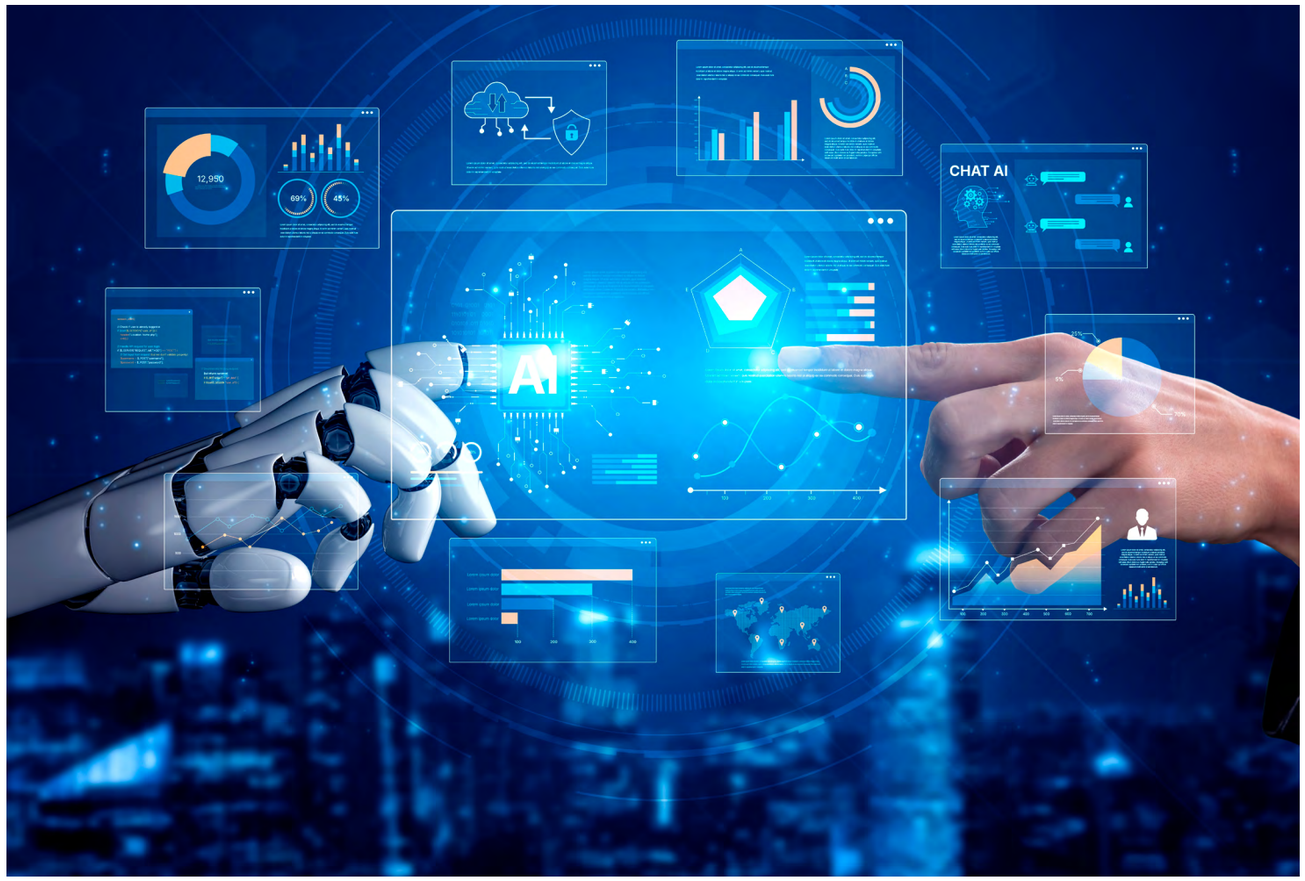
Symantec by Broadcom

Carbon Black. by Broadcom

carahsoft

Building Trust in Mission-Critical AI Systems

As AI takes on greater responsibility across government, agencies must prioritize visibility, reversibility, continuous monitoring and data-centric security.



The Strategic Pivot to Mission-Critical Authority

The federal AI landscape has shifted from experimentation to an operational imperative. While early adoption primarily focused on enhancing administrative productivity — leveraging large language models (LLMs) for document summarization or back-office automation — the federal government is now crossing into the autonomous frontier.

This transition represents a strategic pivot toward mission-critical systems where the stakes are no longer measured in efficiency gains, but in the preservation of national security, the integrity of critical infrastructure and the maintenance of public trust.

Many organizations still operate with disconnected teams, tools, data and environments, creating significant blind spots.

As agencies move beyond basic generative assistants toward autonomous decision-making engines, the underlying infrastructure must be prepared to handle high-stakes environments that directly impact the citizenry.

Unlike commercial entities that may adopt a “pilot first, govern later” mentality, federal agencies handle highly sensitive data domains, including controlled unclassified information (CUI), PII and HIPAA-regulated health records. In this theater, a single failure of security or logic is not a mere budgetary setback; it is a compromise of the national interest.

Federal agencies are moving with extreme urgency, driven by an administration-wide push and an overriding sense that they must adapt or be left behind. To secure this frontier, leadership must look past simple governance and invest in advanced technical guardrails and a phased modernization of the federal IT estate. (ctd.)

Photo Credit: Summit Art Creations/Shutterstock

Securing High-Impact AI: Beyond the Sandbox

As AI integration moves into “high-impact” use cases — defined by the NIST AI RMF as systems affecting public safety, emergency response, civil rights or critical infrastructure — agencies must have a more advanced security approach.

Traditional security controls were designed for predictable software environments. High-impact AI systems require additional safeguards because model outputs can directly influence mission-critical decisions. In these high-stakes environments, the potential for algorithmic drift, data poisoning or intentional logic manipulation requires a shift from permissive sandboxes to a “tollgate” architecture.

Before any high-impact system can transition to a live operational environment, it must survive two rigorous pre-deployment hurdles:

- **Adversarial Red-Teaming:** This involves multi-stage offensive testing in which dedicated teams simulate the tactics of sophisticated threat actors. These teams intentionally attempt to trick the model, exploit its reasoning vulnerabilities or bypass safety guardrails. This is not a “check-the-box” exercise but a stress test of the model’s logic under fire.
- **AI Impact Assessments:** Program managers must produce formal documentation that quantifies model output quality and rigorously weighs mission benefits against potential societal or operational harms. This assessment defines the specific mitigation strategies required to maintain the system’s integrity over its lifecycle.

These testing regimens enable agencies to innovate while maintaining confidence in AI performance, security and compliance. By producing a verifiable audit trail of model quality and risk mitigation, these tollgates provide the necessary evidence for the CAIO to issue a formal authorization to operate.

Under OMB M-25-21, the CAIO bears the responsibility for the system’s

Photo Credit:
Zamrznuti tonovi/Shutterstock

mission outcomes; these pre-deployment protocols allow the CAIO to sign off on deployments with high confidence. Rather than stalling progress, these safeguards transform AI from a liability into a legally defensible and operationally resilient mission asset.

The Cryptographic Shield: PETs*

PET Strategy	Operational Focus	Primary Use Case	Resource Implications
Differential Privacy / Data Obfuscation	Targeted at real-time, low-latency edge inference.	Shielding underlying source records during active operational queries.	Low overhead; optimized for high-velocity mission environments.
Fully Homomorphic Encryption (FHE)	Targeted at high-security, high-intensity offline computation.	Performing complex mathematical analysis on encrypted data without decryption.	Extremely high; reserved for sensitive cases where compute cost is secondary to security.

Security controls must travel with the data across endpoint, network and data environments. By embedding protection within the data fabric, agencies ensure that information remains unreadable and unusable to unauthorized actors, even if a network boundary is breached.

PETs can help address key sovereign AI challenges by allowing agencies to analyze sensitive data while maintaining greater control over security and governance requirements. As agencies evaluate the data exposure risks of multi-tenant public clouds, there is a growing trend toward repatriating workloads to self-hosted, private cloud environments. PETs allow agencies to

*PETs: privacy enhancing technologies

Photo Credit: amgun/ Shutterstock

perform complex analysis on sensitive data within these private enclaves without sacrificing the ability to scale. This architecture ensures that the agency retains greater control over data lineage and integrity, satisfying the most stringent national security requirements while still leveraging the power of advanced algorithmic analysis.

The Tactical Veto: Engineering Human-in-the-Loop and Reversibility

While advanced encryption and automated validation are vital, they do not replace the fundamental necessity of human accountability. In autonomous systems, algorithmic drift — the gradual degradation of a model's performance due to changes in real-world data — or "hallucinations" can lead to outputs that negatively impact public welfare. To prevent these failures from resulting in irreversible harm, human-in-the-loop (HITL) architecture must be engineered as a non-bypassable checkpoint.

The centerpiece of this architecture is the reversibility function. This is not merely a policy guideline; it is a deterministic programmatic capability built directly into the software fabric. It provides human operators with a "tactical veto" that can immediately halt and entirely undo an automated decision the moment an anomaly or unsafe output is detected.

The necessity for reversibility addresses a core truth of modern digital transformation: AI doesn't eliminate the need for governance — it exposes where your governance was already weak.

If an agency cannot reverse an automated decision, it has effectively ceded mission control to a black-box algorithm. Engineering reversibility is the ultimate safeguard for constitutional protections and statutory compliance. It ensures that regardless of how sophisticated the AI becomes, the final authority remains with a human who is legally and ethically accountable for the mission's outcome.

(ctd.)

Photo Credit:HappyEva/ Shutterstock

Operationalizing Machine Identity: Securing Agentic AI with Zero Trust

The most complex evolution in the autonomous frontier is the shift toward agentic AI. This shift represents a significant shift in traditional identity and access management approaches.

When an autonomous agent is accessing a database to summarize law enforcement records or update predictive logistics, a password-based identity becomes obsolete. To secure this environment, agentic AI must be integrated into a matured zero trust architecture through specific architectural requirements:

- **Cryptographic Machine Identities:** Every autonomous agent must be assigned a unique, verifiable identity token. This allows the network to distinguish between a sanctioned agent and a malicious process attempting to masquerade as an authorized service.
- **Metadata-Based, Least-Privilege Controls:** Access must be governed at the data level. This relies on the foundation established in Phase 2 of the modernization framework: the application of FAIR principles. By tagging data with rich metadata, the zero trust architecture can enforce policies that ensure an agent only interacts

“If an agency cannot reverse an automated decision, it has effectively ceded mission control to a black-box algorithm.”

with the specific data sets required for its mission.

- **Logical and Micro-Segmentation:** Agencies must utilize logical segmentation and micro-segmentation to isolate agent activities. This creates a ring-fenced environment where an agent's lateral movement is restricted, even if its identity is compromised.

These controls neutralize the risk of lateral privilege escalation, where a compromised agent could be used to move through a network and access unauthorized data. As these architectures mature, the role of the cybersecurity professional will change. Traditional teams will morph into AI governance teams, shifting their focus from defending infrastructure boundaries to governing data lineage and model integrity. This ensures that autonomous agents act only within their defined parameters, protecting the enterprise from cascading autonomous failures.

The Self-Healing Infrastructure: Continuous Monitoring and Drift Detection

The deployment of an AI model marks the beginning of the day two operational challenge. Maintaining a secure and reliable posture requires operationalizing the "manage" and "measure" functions of the NIST AI RMF.

Federal architects must move toward a self-healing infrastructure that monitors real-time model telemetry to detect degradation before it impacts the mission. To achieve full visibility, an organization must adopt platforms, tools and practices that enable it to monitor its infrastructure, networking, applications, data and systems.

By inspecting traffic at the payload level, agencies can detect and block the transfer of PII or PHI to unvetted public models.

Furthermore, by utilizing DLP and XDR capabilities, defenders act as an "instant replay referee," seeing exactly how data interacts with the environment in real-time. When drift or anomalies are detected, an automated feedback loop

Photo Credit: ART STOCK CREATIVEShutterstock

funnels that telemetry back into retraining and fine-tuning cycles, allowing the system to self-correct.

This continuous monitoring loop is the most effective defense against the risks of shadow AI. But many government agencies are struggling to achieve complete visibility because they have a complicated mix of on-prem, cloud-based and hybrid systems, data and applications.

So where do they begin, and how do they simplify their journey to observability excellence? When agencies provide a superior, sanctioned environment that offers visibility and control, employees have no reason to seek out unvetted public tools. By capturing data interactions in real-time, agencies gain the visibility they currently lack, transforming data from an unmanaged liability into a secured, strategic mission asset.

Successfully defending the autonomous frontier is an architectural challenge, not a procurement task. Federal leadership must recognize that integrating AI into the mission requires more than just purchasing the latest software; it requires a commitment to a phased modernization roadmap that transitions logically from administrative governance to enterprise-scale, autonomous operations.

The agencies that succeed in this new era will be those that prioritize trust, visibility and control alongside their innovations. By adhering to the five-phase modernization framework — from establishing foundational CAIO leadership to the securing of agentic systems under zero trust — federal IT decision-makers can protect national security while delivering on the promise of AI.

Photo Credit: Summit Art Creations/Shutterstock