GovCIO MEDIA & RESEARCH

MARCH 2026

DeepDives

Modernizing Federal RISK MANAGEMENT

INSIDE:

SPONSORED BY

AXIAD FEDERAL

From the editor's desk

Sarah Sybert, Managing Editor

Rethinking Government's Risk Tolerance

Federal agencies are rethinking how risk is defined, measured and managed as artificial intelligence and digital capabilities move from pilot programs into mission-critical operations.

Officials are moving away from traditional, compliance-based risk management. Today's speed of technology requires a different approach. AI systems and modern software need different benchmarks for their performance, particularly in decentralized and high-risk environments.

Now leaders are pointing to new models and standards. The White House's AI Action Plan calls on the National Institute of Standards and Technology to revise its AI Risk Management Framework by easing regulatory burdens to enable innovation.

NIST recently released its Assessing Risks and Impacts of AI framework, or ARIA, which uses a layered approach to evaluate AI behavior. This gives agencies deeper visibility into both technical and contextual risk.

Together, these efforts reflect a broader shift toward real-time risk management as federal agencies face growing pressure to deploy innovative technologies securely, responsibly and at scale.

Table of Contents

Federal AI Efforts Push to Prove Mission Value

Agencies aim to demonstrate mission payoff and elevate flexible, risk-based standards as part of their federal AI strategies.

BY SARAH SYBERT

Federal agencies are deploying artificial intelligence use cases that are directly tied to mission amid easing regulations as part of the White House's directives over the past year.

The AI Action Plan calls on the National Institute of Standards & Technology (NIST) to revise its AI Risk Management Framework to ease some of the regulatory burdens, for example.

The revised version is intended to facilitate that and deploy federal AI use cases that "see a payoff that's commensurate to the investment," said NIST AI and Cybersecurity Researcher Martin Stanley during GovCIO Media & Research's AI Summit in Tysons, Virginia. "It's very important that people trust their AI, or else they won't use it. And then that investment becomes not a good investment."

The value-driven approach to AI development has introduced a higher risk tolerance across government. "The reality is that there's a calculated risk you have to take," said Cloudflare Head of Federal Sales Anish Patel at the event.

This has resonated in efforts like the Pentagon's November tech acquisition strategy to speed up technology development and delivery.

"An 85% solution in the hands of our armed forces today is infinitely better than an unachievable 100% solution endlessly undergoing testing," War Secretary Pete Hegseth said in November.

Photo credit: Invision Events

Both speed and trust are key components to services like the Navy, which is building a hybrid fleet of both manned and unmanned platforms operating in a distributed environment.

"We need to operate without necessarily a whole lot of direct bidirectional comms to the beach, and so, to be able to do this in a somewhat autonomous manner," said Chris Page, deputy director of the Chief of Naval Operations' intelligence division. "There's a point, the earliest time intelligence of value and the latest time intelligence of value, and we need to make sure that our analytical processes … deliver the answer that's needed within the right time frame. And there's a level of accuracy required."

The Marine Corps stood up three digital transformation teams over the past year to deliver emerging technologies to the command and integrate into Marine Corps headquarters to inform policies and standards, said Capt. Christopher Clark, AI lead for the service's Deputy Commandant for Information. The service is currently developing additional teams to launch over the coming year.

"The reward has to be significant to invest in some of this technology. It is very expensive, and so having those teams out there where they're embedded in the mission problem, they understand the problems they face day-to-day, and bringing that information back to headquarters Marine Corps, where we aren't necessarily in that fight, is extremely valuable in understanding where is the value proposition, what is going to make the Marines faster, more lethal, make better decisions," Clark added.

“[Use cases must] see a payoff that’s commensurate to the investment.”

—Martin Stanley, Principal AI and Cybersecurity Researcher, NIST

GovCIO MEDIA & RESEARCH

DeepDives

Digital Identity Risk Management Process Flow

NIST's fourth revision of its Digital Identity Guidelines outlines the processes and technical requirements for digital identity assurance levels. The guidance enables organizations to more effectively address both mission needs and user experience in digital identity technology.

Source: NIST

PARTNER INTERVIEW

AXIAD FEDERAL

Risk in the Age of Quantum

Doing more with less requires careful consideration of technology assets and post-quantum cryptography development.

Andrew Sheedy
Director, Federal Team, Axiad

What are the biggest ICAM and PKI challenges government agencies face today?

Sheedy Many agencies are also running 20-year-old public key infrastructure (PKI) systems on outdated hardware, which leads to higher total cost of ownership from hardware refresh cycles, software licenses and specialized staff.

Their IT staff are often overwhelmed by the need for manual certificate management and derived PIV issuance that runs on manual processes. This can lead to human error resulting in expiration-related outages, so remote users are unable to access critical systems when they need them. OMB M-22-09 requires phishing-resistant MFA agency-wide, but most agencies still rely on basic PIV that doesn't support mobile devices, workers that are not eligible for PIV and work environments that don't support PIV. There are limited options for legacy-derived PIV, and they are expensive and require operationally heavy manual processes. Agencies must also prepare now for Q-day, the post-quantum event

that's looming on the horizon in the next five to seven years. Q-day will occur when quantum computers break current encryption models.

The current legacy PKI systems are unable to support new post-quantum algorithms, and they have no roadmap for when they will support them. All our PKI-dependent systems will have to migrate before the quantum threat materializes, which will require a massive multi-year effort.

Bad actors are currently utilizing "harvest now, decrypt later" attacks which are already collecting encrypted data for future decryption.

What concrete steps can agencies take to overcome these challenges?

Sheedy The first step is to look at ways agencies can offload some of the traditional, on premises, legacy infrastructure to the vendor community. At Axiad, we've developed a FedRAMP moderate cloud-certified PKI and credential management platform. It enables an agency to offload some of the traditional legacy on-premises PKI to a low-touch managed service that is FedRAMP authorized.

Begin with specific populations such as remote workers — mobile users or contractors — then target the applications with the highest security risk or user friction. As a migration strategy, start with less critical systems and move to increasingly critical systems over time. This strategy will demonstrate return on investment (ROI) before they implement an agency-wide rollout.

Make sure your new modern credential management and PKIaaS can integrate with your existing identity providers (IdPs) and legacy Certificate Authorities (CAs), so your migration is a process rather than rip and replace.

Using a migration strategy over time, you should be able to gradually reallocate budget while addressing current security mandates. If you do this, you should be on a modern PKI with support for post-quantum cryptography algorithms by the time Q-day arrives. (ctd.)

"With Q-day potentially only five years away and current austerity measures, modernizing PKI now means one migration instead of two complex migrations. This will save the agency time, budget and resources while addressing multiple critical security requirements simultaneously."

—Andrew Sheedy, Director, Federal Team, Axiad

This is a strategic approach that Axiad supports with its FedRAMP Moderate credential management system with PKIaaS. We include extensive integrations with existing IdPs and CAs to ease the migration process. Our goal is to remove the infrastructure burden while helping agencies meet the FIPS 140-3 requirements by the September 2026 deadline and prepare them for post-quantum cryptography.

They should deploy “self-service derived PIV” which doesn’t require a help desk call. This can extend PIV authentication to mobile devices without physical cards. It also enables users to self-provision their credentials in minutes, not days. That can reduce help desk calls by 50% by eliminating manual requests.

Agencies will have to look for budget opportunities to support cybersecurity modernization in this time of austerity measures. Do a cost analysis to compare hardware security module (HSM) upgrade costs versus cloud PKI subscription ROI. Redirect the hardware budget to an operational expense model that scales with actual usage.

Document that the agency is doing more with less by addressing compliance, quantum readiness and efficiency simultaneously. Self-service derived PIV and certificate provisioning can reduce IT workload by 70-80% compared to manual processes. Automated certificate lifecycle management can eliminate expiration-related outages. By freeing IT staff from routine PKI maintenance, they can be redeployed to focus on strategic initiatives.

Photo Credit: Frame Stock Footage/Shutterstock

What compelling events should agencies be watching out for in the near or medium term as triggers for action?

Sheedy There are immediate compliance deadlines for zero trust implementation milestones such as those established by OMB M22-09. These include implementing phishing resistant authentication for both employees who are covered by PIV and their long-term contractors, but also agency partners — people that are outside of the agency's affiliate boundary.

The FIPS 140-2 cryptographic standard, which has been the time-honored gold standard for crypto, will be phased out in favor of FIPS 140-3 which will take effect in September of 2026.

Agencies should take advantage of critical infrastructure events to drive migration so they can redirect their existing budgets instead of needing additional funding. Treat these as opportunities to migrate to new tech instead of retrofitting existing technologies.

Agencies should also use operational incidents to get the resources they need for migration. These could include certificate expirations causing application outages, successful phishing attacks due to weak authentication and audit findings that cite PKI vulnerabilities or compliance gaps.

The coming quantum event is the most compelling trigger. CISA and NIST have taken a strong leadership role in providing a roadmap. But it's not enough to just read those documents and form your Cryptographic Center of Excellence. You must start taking action to test these things. You need to start small and get some early wins. This enables you to build some momentum and grow stakeholder support inside the agency to help get the necessary funding.

Photo Credit: PeopleImages/Shutterstock

Modernized PKI and Derived PIV. Automate Now.

The only FedRAMP™ Moderate ATO Credential Management with PKIaaS

You know the challenge: Deliver phishing-resistant access for use cases that can't use PIV cards, upgrade HSMs to comply with FIPS 140-3 by September 2026, and prepare for post-quantum cryptography. And do it with fewer people and tighter budgets.

The good news?
Axiad Conductor has already been proven in Federal agencies.

- Credential Lifecycle Automation at Scale
- Derived PIV at Speed
- PQC Readiness on Schedule

Vist us: federal.axiad.com/request-mission-brief

10X

OPERATIONAL EFFICIENCY

Get Compliant.
Reduce Your IT Workload.

FedRAMP™ | AXIAD FEDERAL

NIST's ARIA Adds Red Teaming to AI Evaluation

The framework aims to measure real-world AI behavior, moving beyond accuracy scores to capture risks revealed through model testing, red-teaming and field trials.

BY ROSS GIANFORTUNE

The National Institute of Standards and Technology's new evaluation environment assesses how AI systems behave when used by people in real-world settings.

Assessing Risks and Impacts of AI (ARIA) 0.1, unlike previous NIST evaluations that focused on accuracy, bias or discrete technical capabilities, measures system behavior across real contexts and use cases.

"AI systems are extremely complex," NIST's Information Access Division Chief Mark Przybocki told GovCIO Media & Research in an interview. "A single number to characterize the level of performance is often insufficient."

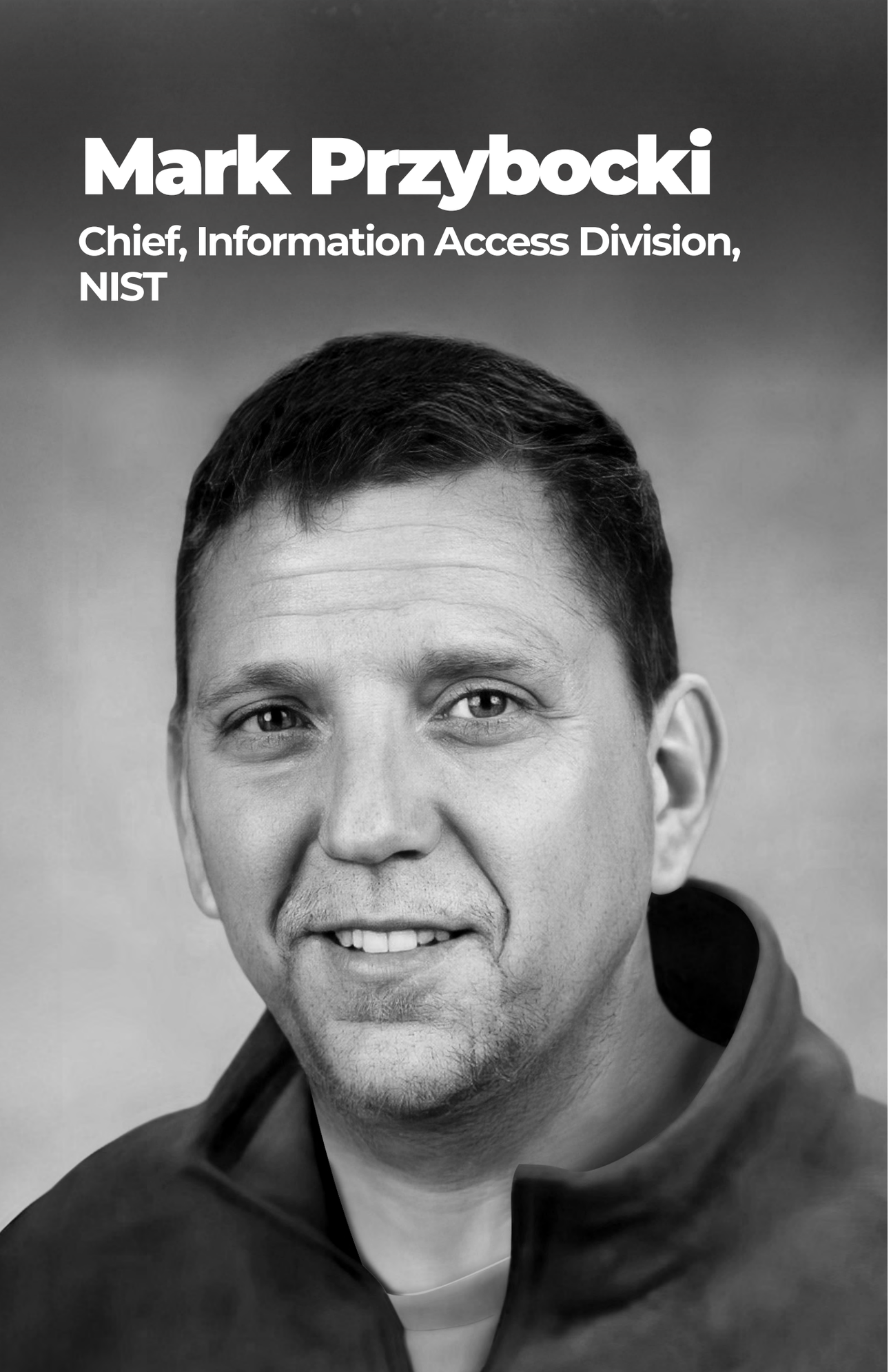
Instead, ARIA uses a three-tier structure — model testing, red-teaming and field testing — to generate a multidimensional view of system behavior.

A Layered Approach

NIST publicly released its first ARIA evaluation plan in December, reflecting what Przybocki described as growing recognition that traditional benchmarks fail to capture the full range of AI risks.

According to the plan, each tier examines different aspects of robustness and risk. Model testing assesses whether a system performs as advertised. Red-teaming probes adversarial or malicious use. Field testing evaluates how

Photo Credit: Digitala World/Shutterstock



systems behave when deployed in realistic scenarios with human users.

Red-teaming is a central component of ARIA's design. While model testing verifies stated capabilities and field testing examines everyday use, red-teaming is intended to uncover unanticipated behaviors.

"By incorporating human testers into AI evaluation," he added, "red-teaming and field testing can help to reveal positive and negative impacts that are not known ahead of time or that relate to AI use in its operating environment."

Such findings can also inform future evaluations.

Measuring 'Technical and Contextual Robustness'

While existing evaluation frameworks often emphasize benchmark scores or accuracy, ARIA assesses how AI systems perform across varied real-world conditions, according to the evaluation plan.

In the pilot, NIST is focusing on large language models (LLMs) "due to the immediate need to understand the wide variety of contexts in which they may be used across private industry and the public sector," Przybocki said.

He added that ARIA's metrics differ from traditional evaluations in several ways. They combine expert annotation with human tester feedback. They also aggregate results into a multi-scale structure that allows evaluators to zoom in on specific behaviors or zoom out for a high-level score.

A Sector-Agnostic Framework with Sector-Specific Flexibility

Although ARIA 0.1 focuses on LLMs, the program is sector-agnostic, Przybocki explained. The goal is to build a framework across government services, critical infrastructure and commercial applications.

"NIST recognizes the importance of both generalizable and context-specific evaluation approaches," he said.

In the pilot, the team selected impacts that could materialize across multiple sectors, such as misinformation, harmful content or flawed reasoning. Future iterations will allow agencies to tailor robustness scores to their own risk tolerances.

The evaluation plan emphasizes that ARIA is not a certification program. Instead, it will help organizations understand how AI systems behave under varied conditions and to inform future standards and best practices.

Why Now

The industry is grappling with how to evaluate AI systems that increasingly shape public life. Traditional benchmarks are often built around static datasets and therefore struggle to capture the dynamic, context-dependent risks of generative AI. Przybocki said that ARIA is meant to fill that gap.

"ARIA seeks to improve AI evaluation by accounting for these varied contexts via realistic scenarios and multiple testers," he said.

"A single number to characterize the level of performance is often insufficient."

—Mark Przybocki, Chief, Information Access Division, NIST