ATARC

WHITEPAPER

Adopting an AI-Enhanced Risk-Based Approach for Cybersecurity Efficiency

IN PARTNERSHIP WITH

Qualys.

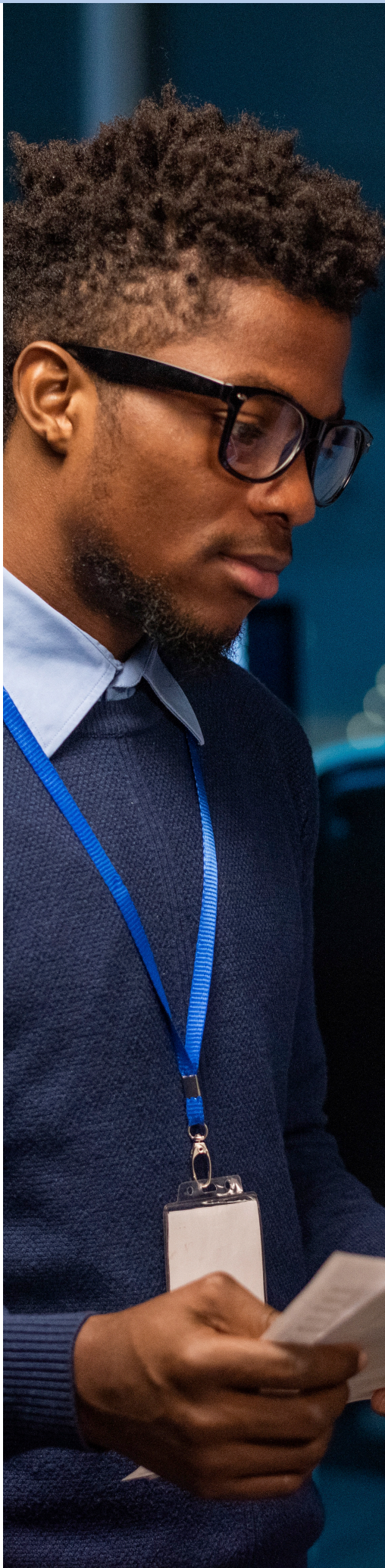
Executive Summary

Federal agencies are entering a new phase of cybersecurity defined by the rapid integration of artificial intelligence into enterprise systems, security tooling, and mission applications. Unlike prior technology shifts, AI introduces a fundamentally different risk model—one characterized by autonomy, opacity, and scale.

Participants described a widening gap between **the speed of AI adoption** and the **maturity of governance, acquisition, and risk frameworks** needed to manage it. Traditional approaches—built around static systems, predictable behavior, and clearly attributable human decision-making—are increasingly insufficient.

Three structural pressures are emerging simultaneously. Agencies are losing visibility into their environments as AI becomes embedded in platforms and workflows. Governance models are fragmenting, with responsibility distributed across multiple offices without clear authority. At the same time, risk is becoming harder to define, as existing frameworks were not designed to account for autonomous or probabilistic systems.

Despite these challenges, agencies face mounting pressure to deploy AI capabilities in support of mission outcomes. The result is a persistent tension between operational urgency and risk discipline—one that is unlikely to resolve without deliberate structural change.



1. The Shift to AI-Driven Cyber Risk Models

Cybersecurity in the federal enterprise has traditionally focused on protecting infrastructure, networks, and user identities. AI disrupts this model by introducing systems that are not strictly deterministic. Their behavior can vary based on training data, environmental inputs, and internal optimization processes.

This creates a new category of risk that is less about system compromise and more about **system behavior under uncertainty**. Agencies are no longer securing static assets alone; they are increasingly responsible for systems that generate outputs, make recommendations, and in some cases act autonomously.

As a result, cybersecurity is evolving from a control-based discipline into one that must incorporate probabilistic reasoning. Risk can no longer be fully enumerated in advance, and outcomes cannot always be predicted through traditional testing. This shift requires a reassessment of how agencies define assurance and accountability.

2. The Visibility Problem: Assets, Agents, and Unknowns

A central concern raised throughout the discussion was the erosion of visibility. Agencies have long struggled with asset inventory, but AI introduces new layers of complexity that make the problem more acute.

AI capabilities are increasingly embedded within commercial platforms, often without clear delineation of where or how they operate. At the same time, users and developers are beginning to create or deploy autonomous agents that interact with enterprise systems in ways that are not centrally tracked.

The result is an environment in which agencies may not fully understand what systems are operating, what data they are accessing, or what actions they are capable of taking. This lack of visibility extends beyond infrastructure to include decision-making processes themselves.

Without a clear understanding of what exists in the environment, risk management becomes largely reactive. The concern is not simply that unknown systems exist, but that those systems may be acting with a degree of autonomy that was not previously possible.

3. Governance Breakdown: CIO, CDAO, and Emerging AI Authorities

The introduction of AI has disrupted established governance models in federal agencies. Responsibilities that were once clearly aligned under the CIO are now distributed across multiple entities, including Chief Data Officers, privacy offices, and newly established AI leadership roles.

While this reflects the cross-cutting nature of AI, it has also created ambiguity around decision authority. In practice, agencies are navigating overlapping review processes, unclear escalation paths, and inconsistent standards for approval.

This fragmentation is particularly evident in cases where systems have been developed but cannot be deployed due to unresolved governance questions. The issue is not a lack of oversight, but rather a lack of coordinated oversight with clearly defined roles.

At the same time, there is an inherent tension between the need for rigorous review and the demand for rapid capability delivery. As AI becomes more central to mission execution, delays in decision-making can have operational consequences. Resolving this tension will require more than new offices; it will require clearer alignment of authority and accountability.

4. The Rise of Autonomous Agents and Non-Human Identities

One of the most consequential developments discussed was the emergence of AI agents as operational actors within federal environments. These agents are capable of executing tasks, interacting with systems, and in some cases making decisions with limited human intervention.

This introduces a new class of identity—entities that are not human, but that act on behalf of humans or organizations. While similar in some respects to service accounts, these agents differ in their potential for autonomy and variability. The challenge is not simply technical, but conceptual. Agencies must determine how to grant access, assign responsibility, and monitor behavior for entities that do not fit traditional identity models. A useful framing that emerged from the discussion is to treat these agents as analogous to junior personnel: capable of performing useful work, but requiring supervision, constrained authority, and progressive trust.

This approach highlights the need for governance models that account for both capability and maturity, rather than assuming uniform reliability across all automated processes.

5. Risk Without Clarity: Embedded AI in COTS and SaaS

The increasing prevalence of AI within commercial platforms presents a distinct and immediate challenge. Agencies are acquiring systems that include embedded AI capabilities, often without full transparency into how those capabilities function.

In many cases, the use of AI cannot be easily separated from the core functionality of the product. Disabling it may degrade performance or eliminate key features, effectively forcing agencies to accept the associated risks.

These risks are not fully addressed by existing frameworks. While programs such as FedRAMP provide important baseline assurances, they were not designed to evaluate the behavior of AI systems, the provenance of training data, or the potential for unintended outputs.

Compounding this issue is the shared nature of many SaaS environments, where multiple tenants rely on the same underlying infrastructure. This creates potential for systemic impact if issues arise, as well as dependencies on vendor controls that agencies do not directly manage.

The result is a growing reliance on external assurances in areas where internal validation remains limited.

6. Acquisition Friction and Evaluation Gaps

AI is also introducing new complexities into the acquisition process. Agencies are increasingly faced with multiple solutions that offer similar capabilities but rely on different underlying AI approaches.

Evaluating these options requires expertise that is not always present within traditional acquisition or technical evaluation teams. As a result, decision-making may focus on functionality while underestimating differences in risk profiles.

In addition, governance requirements may necessitate separate review of AI components, further extending timelines and increasing uncertainty. This can lead to situations where viable solutions are delayed or abandoned due to the difficulty of navigating the evaluation process.

There is a clear need for more structured approaches to assessing AI within acquisition, including standardized criteria and integrated evaluation processes.

7. Human Judgment in an AI-Augmented Environment

The role of the human operator remains central, but it is evolving in ways that introduce new risks. Participants noted that users may increasingly rely on AI-generated outputs without fully understanding their limitations.

This dynamic is particularly concerning in environments where accuracy is critical. Errors introduced by AI—whether through hallucination, bias, or misinterpretation—may not be immediately apparent, especially if the output appears authoritative.

Maintaining effective human oversight requires more than simply keeping a person "in the loop." It requires ensuring that individuals have the skills, confidence, and institutional support to question and validate AI outputs.

This represents a cultural as well as a technical challenge. Organizations must reinforce the expectation that AI is a tool to be interrogated, not an authority to be accepted without scrutiny.

8. Workforce Transformation: From Individuals to Augmented Teams

The integration of AI into daily workflows is beginning to reshape expectations of the federal workforce. Increasingly, effectiveness is tied not only to individual expertise, but to the ability to leverage AI tools in support of that expertise.

This shift suggests a future in which personnel operate as part of augmented teams, combining human judgment with machine-driven capabilities. In some cases, individuals may be expected to manage multiple AI agents as extensions of their own work.

While this model offers clear productivity benefits, it also introduces new dependencies and potential vulnerabilities. The effectiveness of the workforce becomes partially contingent on the reliability and security of the AI systems it employs.

Developing the necessary skill sets will require a deliberate focus on both technical literacy and critical thinking, ensuring that personnel can effectively integrate AI into their work without becoming overly reliant on it.

9. Training and Scenario-Based Readiness

Given the complexity of AI-driven environments, traditional training approaches are insufficient. Participants emphasized the value of immersive, scenario-based exercises that simulate real-world decision-making under uncertainty.

These approaches allow individuals to experience the consequences of AI-related decisions, understand trade-offs, and develop the judgment required to operate effectively in dynamic environments. They also provide a mechanism for testing organizational processes and identifying gaps in governance or coordination.

Such training models, long used in military contexts, may be increasingly relevant for civilian agencies as they navigate the challenges of AI integration.

---

10. Strategic Considerations for Federal Agencies

The discussion points to several strategic priorities that agencies should consider as they adapt to AI-enabled environments.

Agencies must first establish a more complete understanding of where and how AI is being used within their environments. This includes not only systems explicitly identified as AI-enabled, but also capabilities embedded within broader platforms.

Governance structures should be clarified to ensure that decision authority is both well-defined and aligned with mission needs. This may involve consolidating responsibilities or creating clearer integration points between existing roles.

Risk frameworks will need to evolve to address the unique characteristics of AI, including its dependence on data, its potential for autonomous behavior, and its lack of transparency. Existing models provide a foundation, but they require extension rather than simple application.

Finally, agencies must invest in the workforce, not only by building technical skills, but by reinforcing the critical thinking and judgment required to operate effectively in AI-augmented environments.

Conclusion

The integration of AI into federal cybersecurity is not a future challenge; it is a present reality. Agencies are already operating in environments where visibility is incomplete, governance is evolving, and risk is not fully understood.

The path forward is not to slow adoption, but to align structures, processes, and capabilities with the nature of the technology itself. This will require a shift in how agencies think about systems, identities, and decision-making.

Ultimately, success will depend on the ability to manage not just technology, but **intelligent, adaptive systems** operating within complex and mission-critical environments.