



# Beyond the Hype: Turning AI into Mission Advantage in Federal Government

A Chatham House Rules  
Roundtable White Paper  
*(Not for Attribution)*

IN PARTNERSHIP WITH



## Executive Summary

Artificial intelligence has moved rapidly from experimentation to expectation across the federal government. Leaders are no longer asking whether to adopt AI, but how quickly it can be integrated into mission operations. Yet beneath this urgency lies a more complicated reality: while tools have advanced dramatically, the underlying conditions required to make AI effective—data maturity, governance, cost discipline, and workforce readiness—have not kept pace.

A recent cross-agency roundtable of federal practitioners revealed a consistent theme: **AI is not the hard part.** The real challenge lies in operationalizing it within environments defined by fragmented data, complex policy constraints, and evolving security requirements. Agencies are simultaneously trying to modernize infrastructure, retrain their workforce, and deploy emerging capabilities—all while managing risk in high-stakes mission environments.

What emerges is not a story of failure, but of friction. Progress is happening, but incrementally. As one participant put it, “we’re moving forward—but it feels like ten meters at a time.”

This paper explores that friction and outlines a path forward for federal leaders seeking to translate AI enthusiasm into sustained mission advantage.

## The Illusion of AI Readiness

There is a growing perception among senior leaders that large language models represent an “easy button” for transformation. The promise is compelling: faster analysis, automated workflows, and immediate insights from vast data stores. But practitioners working at the operational level describe a different reality—one where the introduction of AI often exposes deeper systemic issues rather than resolving them.

In environments such as law enforcement, defense, and oversight, agencies are already overwhelmed by the sheer volume of data they collect. The challenge is not access, but usability. Massive datasets—ranging from cyber telemetry to procurement records—exist across disconnected systems, governed by different rules and often inaccessible to those who need them most. AI, in this context, becomes another layer of complexity.

One participant captured the dynamic succinctly: the organization is “sitting on more data than we can possibly process,” and AI is being introduced into that environment without first resolving how the data should be structured, shared, or governed.

The result is a growing gap between expectation and execution. Leaders expect rapid transformation; operators are still trying to make the data usable.



## Data: The Enduring Constraint

If there was a single point of consensus across the discussion, it was this: **data remains the central limiting factor** in federal AI adoption. Not because it is scarce, but because it is fragmented, inconsistent, and often poorly contextualized.

Much of this fragmentation is by design. Legal authorities, classification levels, and mission-specific requirements create natural barriers to data sharing. In some cases, identical or overlapping datasets exist across agencies but cannot be reconciled due to policy or governance constraints. Even within a single organization, different mission areas may operate on entirely separate data environments.

This fragmentation becomes particularly problematic when agencies attempt to apply AI at scale. Modern AI systems rely heavily on processes such as chunking, embedding, and retrieval. When these processes are poorly executed—as they often are in early implementations—they degrade the quality of outputs in ways that are not immediately obvious to end users.

One technical expert noted that many organizations assume their AI systems are underperforming due to limitations in the models themselves. In reality, the issue often lies in how the data was prepared. “It’s not that the model is bad,” he explained. “It’s that the data was chunked and embedded in a way that lost context.”

This loss of context is more than a technical nuance—it directly impacts mission outcomes. In investigative or operational environments, incomplete or misinterpreted information can lead to missed threats, flawed analysis, or incorrect decisions.

Compounding the issue is the pace of technological change. Methods for structuring and enriching data are evolving rapidly, meaning that data ingested into AI systems today may need to be reprocessed within months to remain useful. Agencies that treat data ingestion as a one-time activity risk locking themselves into outdated architectures.

## Governance in a Time of Acceleration

While data challenges are foundational, governance represents an equally significant barrier to progress. Federal agencies operate within a dense web of policies, compliance requirements, and oversight mechanisms. These structures are essential for accountability, but they are not designed for the pace of change introduced by AI.

Participants described a recurring pattern: innovation initiatives stall not because they are prohibited, but because it is unclear whether they are allowed. Over time, assumptions about policy harden into perceived constraints, even when no formal restriction exists. This creates a culture of hesitation, where teams default to inaction rather than risk non-compliance.

At the same time, governance processes themselves are struggling to adapt. Procurement cycles remain slow, even as technology evolves on a monthly cadence. Oversight organizations face the additional challenge of adopting AI tools while also being responsible for auditing their use across government.

There is a growing recognition that governance must evolve from a model centered on restriction to one focused on enablement. This does not mean reducing oversight, but rather clarifying it—establishing clear boundaries within which innovation can occur safely and confidently.



## The Coming Cost Reckoning

Another theme that emerged with increasing urgency is cost. Many agencies are currently benefiting from low-cost or pilot-phase access to AI capabilities, creating the impression that these tools can be scaled with minimal financial impact. Practitioners, however, see a different trajectory.

Large language models are inherently compute-intensive, and their costs scale with usage. When multiple users repeatedly perform similar queries, agencies effectively pay for the same computation multiple times. Data movement—particularly in cloud environments—introduces additional costs, including egress fees that are often overlooked in initial planning.

One participant warned that agencies are becoming “addicted” to inexpensive access, without fully understanding the long-term financial implications. As usage expands, these costs are likely to surface quickly, potentially outpacing budgets that were not designed for this level of demand.

This challenge is not purely financial—it is also strategic. Without clear guidance on when and how to use AI, agencies risk applying expensive tools to problems that could be solved more efficiently through traditional analytics or automation.

## Architecture at a Crossroads

The discussion also highlighted a fundamental shift in how federal systems are being designed. The long-standing “cloud-first” approach is giving way to a more nuanced reality in which hybrid architectures are not just preferred, but necessary.

Sensitive data, particularly in defense and intelligence environments, often cannot be moved to the cloud. At the same time, the computational demands of AI frequently exceed what can be supported on-premises. Agencies are therefore forced to operate across both environments, making deliberate decisions about where data resides and where processing occurs.

This complexity is compounded by the speed of technological change. New approaches to data processing, model design, and system integration are emerging rapidly, raising concerns about long-term sustainability. Several participants cautioned against building rigid architectures around current technologies, noting that many of today’s approaches may be obsolete within a year.

At the same time, there is growing recognition that AI may fundamentally reshape the application landscape. Instead of relying on a proliferation of purpose-built systems, agencies may increasingly use AI-driven interfaces layered over shared data environments. In this model, the distinction between “applications” and “data” begins to blur, potentially reducing the need for many legacy systems.

## The Human Factor

Despite the technical complexity of these challenges, participants consistently returned to one point: **the hardest problem is not technology—it is people.**

End users are not asking for sophisticated tools; they are asking for clarity. A contracting officer does not want to analyze a spreadsheet—they want to know which transactions require attention. An analyst does not want more data—they want better insight.

At the same time, many users are deeply accustomed to existing workflows. Shifting from manual analysis to AI-assisted decision-making requires not just new tools, but new ways of thinking. In some cases, this represents a complete inversion of how work has traditionally been performed.

Training, therefore, becomes critical—not just in how to use AI, but in how to trust it, question it, and integrate it into decision-making processes. Without this transformation, even the most advanced systems will fail to deliver meaningful impact.

## Security, Privacy, and Trust

No discussion of federal AI adoption is complete without addressing security and privacy. These concerns are not theoretical—they are central to how agencies operate.

Participants raised concerns about the risk of exposing sensitive data through AI systems, particularly when models aggregate information in ways that may reveal unintended insights. In highly classified environments, even the combination of seemingly benign data points can create new security risks.

To mitigate these concerns, some organizations are turning to synthetic data for training and experimentation. Others are implementing strict controls on what data can be used and how models are monitored over time.

There is also growing awareness of the need for continuous oversight. AI systems are not static; they evolve, drift, and change in response to new inputs. Ensuring their reliability requires ongoing monitoring and adjustment—an operational burden that agencies are only beginning to fully understand.

## Conclusion: From Momentum to Maturity

The federal government is not standing still. Agencies are experimenting, learning, and gradually advancing their AI capabilities. But the transition from experimentation to operational impact will require a shift in focus.

The path forward is not about adopting the latest model or tool. It is about building the conditions under which those tools can succeed: high-quality data, clear governance, sustainable architectures, and a workforce prepared to operate in a new paradigm.

What emerged most clearly from this discussion is that progress is not linear. It is iterative, uneven, and often constrained by factors outside the control of individual teams. Yet there is also a sense of inevitability. The direction is clear, even if the pace is slow.

AI will not replace the foundational work required to modernize federal systems. But for agencies willing to confront those challenges directly, it offers something far more valuable: the opportunity to transform how decisions are made, how missions are executed, and how government delivers on its responsibilities.

