



INSIGHTS FROM A FEDERAL ROUNDTABLE

Accelerating Federal AI Readiness

A Data-Centric Blueprint for Mission Outcomes

 **ATARC** IN PARTNERSHIP WITH
 **NetApp®**

Executive Summary

Artificial intelligence has rapidly moved from experimentation to expectation across the federal government. Agency leaders are being asked to deliver faster decisions, improved mission outcomes, and greater operational efficiency—all powered by AI. Yet despite this urgency, progress remains uneven.

The core issue is not technological capability. Federal agencies have access to advanced models, scalable compute, and a growing ecosystem of tools. The constraint is more fundamental: **the condition of federal data environments**.

Insights gathered from a recent federal roundtable reveal a consistent pattern. Agencies are attempting to operationalize AI on top of fragmented, siloed, and inconsistently governed data. These challenges—longstanding in federal IT—are now being exposed and amplified by AI.

Participants repeatedly emphasized a central truth:
AI success is directly proportional to data maturity.

To move forward, agencies must shift from treating AI as a standalone capability to recognizing it as the outcome of a **data-centric enterprise strategy**. This requires not only modern infrastructure, but also disciplined approaches to metadata, governance, security, and workforce enablement.



The Illusion of AI Readiness

Across government, AI has become synonymous with modernization. Leadership directives emphasize speed—deploy AI quickly, automate processes, and deliver measurable results. However, this push often creates a false sense of readiness.

In practice, AI does not mask underlying issues. It reveals them.

Agencies attempting to deploy AI frequently encounter challenges tied to:

- Poor data quality
- Lack of standardization
- Limited visibility into data assets
- Fragmented ownership across organizations

What emerges is a disconnect between strategic intent and operational capability. While leadership expectations continue to rise, the foundational work required to support AI has not kept pace.

This dynamic creates a cycle where agencies pilot AI successfully in isolated environments but struggle to scale those successes across the enterprise.

Data Fragmentation as the Primary Barrier

Federal data environments are inherently complex. Over decades, agencies have accumulated systems designed for specific missions, programs, and compliance requirements. The result is a landscape where data is:

- Distributed across multiple clouds and on-premises systems
- Stored in diverse formats and schemas
- Owned and managed by different organizational units
- Generated at increasing volume from edge and field operations

Participants described a paradox: agencies possess vast amounts of data, yet lack the ability to effectively use it.

This fragmentation manifests in several ways. Data duplication is common, with similar datasets stored in multiple systems without synchronization. Inconsistent formats make integration difficult, requiring extensive transformation before data can be used. Organizational silos further complicate access, with different teams maintaining separate standards and processes.

Without a unified approach to managing this complexity, agencies face significant barriers to AI adoption. Models trained on incomplete or inconsistent data produce unreliable outputs. Analytical insights remain localized rather than enterprise-wide. Governance becomes difficult to enforce.

The Shift to Data-Centric Architecture

A key theme emerging from the discussion is the need to move beyond infrastructure-centric thinking. Historically, federal IT has focused on systems—applications, servers, networks. AI, however, shifts the focus to data itself.

A data-centric approach prioritizes:

- Visibility into data across environments
- Consistent classification and governance
- Secure and efficient access
- Alignment between data structure and mission needs

This shift is particularly important in hybrid environments, where data resides across cloud platforms, on-premises systems, and edge locations. Rather than forcing all data into a single repository, agencies are increasingly adopting approaches that allow data to remain distributed while still being discoverable and usable.

This reflects a growing recognition of **data gravity**. Moving large datasets is often impractical due to scale, cost, and latency. Instead, agencies are exploring architectures that bring computation to the data, enabling analysis and AI processing without requiring wholesale data movement.

Metadata as the Foundation of AI

If data is the fuel for AI, metadata is the map.

Metadata provides the context necessary to understand and use data effectively. It defines what data represents, how it can be used, and who can access it. Without this context, even well-structured data becomes difficult to operationalize.

Participants identified metadata as one of the most significant gaps in current federal data strategies. Challenges include inconsistent tagging practices, reliance on manual processes, and difficulty maintaining metadata over time.

Two primary approaches to metadata management were discussed: maintaining metadata in external systems such as catalogs, or embedding metadata directly within data assets. Each approach offers advantages, but both require consistent standards and ongoing maintenance.

A key insight is that metadata cannot remain static. As organizations evolve, data must be reclassified and recontextualized. Emerging approaches leverage AI to automate this process, continuously analyzing and updating metadata to reflect current conditions.

Preparing Data for AI: The Unseen Effort

Much of the work required to enable AI occurs before a model is ever deployed. Data must be transformed into formats that are accessible, structured, and meaningful for machine learning systems.

This preparation involves:

- Converting data into machine-readable formats
- Preserving context during transformation
- Segmenting data in ways that align with AI processing

Participants emphasized that seemingly small decisions—such as how data is segmented—can have significant impacts on AI performance. Improper structuring can lead to loss of context, resulting in inaccurate or misleading outputs.

To address this, agencies are investing in more sophisticated data pipelines that incorporate preprocessing, normalization, and validation steps. These pipelines are designed to ensure that data entering AI systems is both high-quality and contextually intact.

Security, Governance, and Risk

AI introduces new dimensions of risk, particularly in environments handling sensitive or regulated data. Federal agencies must balance the need for innovation with strict requirements for security and compliance.

Participants highlighted several areas of concern. Inconsistent data classification makes it difficult to enforce access controls. Users may inadvertently expose sensitive data by inputting it into external AI tools. Additionally, the use of AI-generated outputs in official processes raises questions about accountability and traceability.

Addressing these risks requires integrating security and governance directly into data architectures. This includes applying zero trust principles, implementing role-based access controls, and automating data protection measures such as redaction and monitoring.

Importantly, governance must extend beyond technology to include clear policies and guidance for AI usage. Without this, agencies risk both operational and legal challenges.

The Human Factor: Adoption and Behavior

While much of the focus on AI centers on technology, the role of the workforce is equally critical. Participants described a complex landscape where different parts of the organization are at varying levels of readiness.

Some users lack familiarity with AI and require foundational training. Others are adopting AI tools rapidly, sometimes without fully understanding the associated risks.

This creates a dual challenge: enabling adoption while ensuring responsible use.

Effective strategies focus on aligning AI with individual and mission needs. When users understand how AI improves their work, adoption becomes more natural. At the same time, clear guidelines and training are necessary to prevent misuse.

Communication plays a key role. Agencies that invest in explaining not just how AI works, but why it matters, are more likely to achieve sustained adoption.

Modernizing Without Starting Over

One of the most important insights from the discussion is that agencies do not need to start from scratch. While legacy systems present challenges, they also represent significant investments and institutional knowledge.

Rather than replacing these systems entirely, agencies are finding success by layering modern data capabilities on top of existing infrastructure. This approach allows for incremental progress while minimizing disruption.

By creating shared data environments and enabling interoperability, agencies can bridge gaps between systems and unlock new capabilities without wholesale transformation.



A Path Forward

The path to AI readiness is not a single initiative, but a coordinated effort across multiple domains. It begins with understanding the current state of data and establishing visibility across the enterprise. From there, agencies must implement governance frameworks, modernize architectures, and prepare data for AI applications.

Equally important is investing in the workforce, ensuring that employees have both the skills and the understanding needed to use AI effectively.

Throughout this process, the focus must remain on mission outcomes. AI is not an end in itself, but a means to improve how agencies operate and deliver value.

Conclusion: From Ambition to Execution

Federal agencies are at a pivotal moment. The potential of AI is clear, but realizing that potential requires a shift in approach.

Success will not be determined by access to technology alone. It will depend on the ability to manage data as a strategic asset—one that is visible, governed, secure, and aligned with mission needs.

As agencies move forward, the distinction between AI strategy and data strategy will continue to blur. In practice, they are one and the same.

Those that recognize this—and act accordingly—will be best positioned to move from experimentation to execution, delivering meaningful impact through AI.

