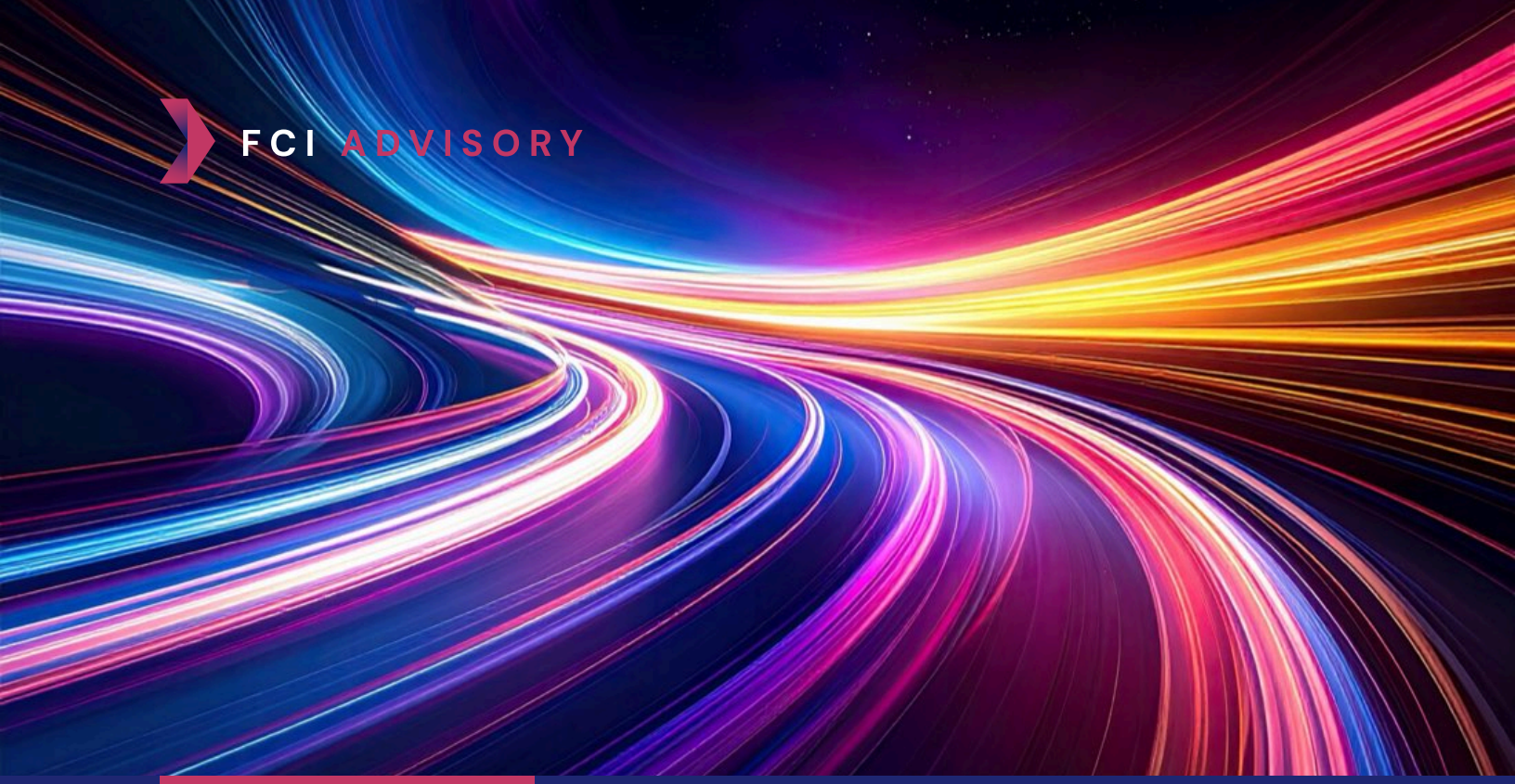

FCI ADVISORY

FEDERAL AI & EXECUTION PRACTICE · RESEARCH BRIEF

Agentic AI in Workforce Systems

Where agentic AI is landing first in federal agencies, the integration realities underneath the deployments, and what program leaders should actually fund.

FCI Advisory · McLean, Virginia

2026

Table of contents

Executive summary

Federal AI coverage tracks two surfaces — the policy layer (OMB memoranda, the NIST framework, executive orders) and the model layer (which foundation model an agency selected, which chatbot landed in pilot). A third layer is doing more of the work and getting less of the attention. Federal agencies are deploying their most ambitious agentic AI implementations inside their own workforce systems first — not customer-facing operations, not analytical back-office work.

~48%
of agentic deployments are workforce / HR — vs ~5% of the press coverage[4]

~5%
share of press coverage workforce / HR AI receives — the inverse of where it lands[1]

6
layers between a foundation model and a deployable federal workforce AI

3
independently sufficient reasons the workforce is the first agentic frontier

This brief sets out why workforce systems are the first agentic frontier in the federal estate, what makes them the hardest agentic implementations anywhere, and what program leaders should fund as a result. Its argument is in four parts:

- **Workforce is where the deployments are.** Coverage focuses on customer-facing and policy AI; deployments concentrate in workforce and HR systems. The category dominating the press is the category where federal agentic deployment is least concentrated.
- **Workforce-first is the right sequence, for three reasons.** Blast-radius asymmetry, a lighter policy burden under M-24-10's administrative category, and a higher marginal-value calculation than commercial workforce AI.
- **The model is the easy part.** The constraint stack underneath — unionized workforce rules, integration, content management, data quality, records governance — is where federal workforce AI actually gets built.
- **Vendor evaluation is scoring the wrong layer.** Programs that score model benchmarks are evaluating the layer where pilots rarely fail. The differentiators are integration depth and content governance.

> *"The model is the easy part. The integration layer, the content layer, the data quality layer, the records governance layer — the stack underneath the model is where federal workforce AI actually gets built. Most attention sits on the wrong layer."*

A note on figures. Use-case shares and directional patterns in this brief are drawn from federal AI use-case inventories and FCI Advisory's engagement base across federal workforce-AI and integration programs, FY24-Q4 through FY26-Q1.[4] They are illustrative of the directional pattern, not point-precise extracts.

01 Where federal agentic AI is actually being built

The federal AI conversation sits in two visible layers. The policy layer above — OMB Memorandum M-24-10,[2] the NIST AI Risk Management Framework,[3] the executive-order architecture. The model layer below — which foundation model the agency chose, which vendor's chatbot is in pilot. Both layers are real and both deserve the coverage they receive.

The third layer is where federal AI actually runs when it runs. It is the integration and content layer — the iPaaS-class middleware that lets an agent reach across federal systems, the Documentum and equivalent content environments that hold the records the agent retrieves from, the data-quality tooling that determines whether the agent's outputs are trustworthy, and the records governance that decides which of the agent's outputs are themselves federal records. This is where most federal AI engineering effort is spent. It is also where almost no public attention lands.

The deployments now going into production are agentic in the technical sense — systems that reason, take actions, integrate across legacy environments, and operate with limited human-in-the-loop. The use cases are workforce-internal: scheduling optimization in unionized operations, position management and classification, employee benefits administration, grievance routing and triage, training and certification pathways, employee-services automation. The programs are large, multi-year, and being awarded right now.

THE MISMATCH: WHERE COVERAGE POINTS VS. WHERE DEPLOYMENTS ARE (ILLUSTRATIVE)

Use-case category	Share of press coverage	Share of agentic deployments
Workforce / HR	~5%	~48%
Analytical / back-office	~15%	~30%
Customer-facing	~50%	~22%
Policy & governance	~30%	—

Share of federal AI press coverage vs. share of federal agentic AI deployments in production, by use-case category. The category dominating the press — customer-facing — is where deployment is least concentrated; workforce / HR is the inverse. Illustrative directional shares aggregated across federal AI use-case inventories and FCI's engagement observation.[14]

This pattern is most visible across the federal infrastructure sectors — postal, surface transportation, and adjacent infrastructure — because those sectors have the largest unionized federal workforces and the most acute operational pressure on workforce systems. But the pattern is a federal technology pattern, not a sector phenomenon. Federal civilian and defense agencies are starting to follow the same path. Workforce-first is counterintuitive against the commercial enterprise AI pattern — where the first agentic deployments land at the customer edge — but in the federal environment it is consistent enough now to call a pattern.

SECTION 02 · THE LOGIC

Why the workforce is the first agentic frontier 02

02 Why the workforce is the first agentic frontier

Three reasons, each independently sufficient.

Reason one: blast-radius asymmetry

FIGURE 02 · HOW FAR AN AI ERROR TRAVELS

Workforce errors stay inside the organization. Customer-facing errors hit the front page.

Federal agencies are choosing the lower-blast-radius use cases first. This is the right sequence; commercial enterprise AI got it wrong.

FCI Advisory framework. The point is the ordering of consequence scope, not precise scaling.

In a federal operational system serving the public, an AI error has immediate external consequences — mis-routed transactions, scheduling failures, service disruptions that show up on the front page. An AI error in a workforce-internal system is recoverable: a grievance gets routed to the wrong queue, a benefits eligibility query escalates for human review, a position-classification draft gets corrected before it ships. The blast radius is contained inside the organization, not externalized to the public. Federal agencies are choosing the lower-blast-radius use cases for their early agentic deployments. This is the right sequence; commercial enterprise AI got it wrong.

HOW FAR AN AI ERROR TRAVELS, BY USE-CASE CATEGORY

Category	Blast radius	What an error looks like
Workforce / HR	Contained	Mis-routed grievance, escalated eligibility query, draft classification corrected before ship. Internal recovery.

Category	Blast radius	What an error looks like
Analytical / back-office	**Operational**	Mis-classified data, bad report, decision-support output trusted incorrectly. Costly to unwind.
Customer-facing	**Public**	Mis-routed transaction, scheduling failure, public service disruption. External, front-page risk.

FCI Advisory framework. The point is the ordering of consequence scope, not precise scaling.

Reason two: a lighter policy burden

OMB Memorandum M-24-10 distinguishes between "rights-and-safety" AI — uses that affect public rights or safety — and administrative AI applications.[2] The rights-and-safety category carries documentation, transparency, and governance requirements that the administrative category does not. Most workforce/HR applications fall in the administrative category. Agencies and vendors building these systems can move faster on workforce AI than on operational AI without leaving the policy framework. The faster path is the path being taken — and it is being taken deliberately, not by accident.

Reason three: value asymmetry

Federal workforce systems are uniquely tangled. They operate under collective bargaining agreements, federal employment law, sovereign-scale continuity requirements, decades of accrued policy precedent, and union-grievance procedures that no commercial HR system has ever had to absorb. Manual processing across these constraints is expensive and slow. An agentic system that can navigate the constraint structure adds value that a commercial-equivalent system would not, because the commercial equivalent does not face the constraint structure. The marginal-value calculation for federal workforce AI comes out higher than for commercial workforce AI, and that math is what is funding the deployments.

03 The constraint stack is the engineering

The agentic systems being deployed in federal workforce functions operate under conditions that make them, plainly, harder engineering problems than most enterprise AI deployments anywhere. The model layer is the easy part. The constraint stack underneath is where the real work is.

FIGURE 03 · SIX LAYERS UNDER THE MODEL

The model is the easiest of the six layers. The constraint stack is where it gets built.

Three layers — integration, content management, records governance — are where the hardest federal-specific engineering lives, and where most attention does not.

Labor-relations	Every agent action must reconcile with collective-bargaining agreements and grievance procedures.	BINDING
Integration / middleware	The agent must reach legacy HR, finance, and case-management systems of record at runtime.	HARDEST
Content management	The records environment the agent retrieves from and writes back to — typically Documentum-class ECM.	HARDEST
Data quality	Source quality that determines whether the agent's outputs are trustworthy or confidently wrong.	HIGH
Records governance	Disposition of the agent's own outputs, which are themselves federal records.	HARDEST
Foundation model	The reasoning engine — the most discussed layer and the easiest to source.	EASIEST

Six layers between a foundation model and a deployable federal workforce AI. FCI Advisory framework, sized to the federal workforce-AI deployment pattern.

Unionized workforce constraints come first

Every action an agent takes — every recommendation, every routing decision, every escalation — is potentially subject to grievance procedures negotiated decades ago and never contemplated with agentic intermediaries in mind. A grievance filed against an agent's decision creates a question no commercial AI deployment has had to answer at scale: who is responsible for the decision, the model or the operator? The agencies and the unions are answering this question in real time. The answers being negotiated now will set precedent that propagates outward.

SIX LAYERS BETWEEN A FOUNDATION MODEL AND A DEPLOYABLE FEDERAL WORKFORCE AI

Layer	What it requires
Labor-relations layer	Every agent action must be reconcilable with collective-bargaining agreements and grievance procedures.
Integration / middleware	The agent must reach legacy HR, finance, and case-management systems of record at runtime.
Content management	The records environment the agent retrieves from and writes back to — typically Documentum-class ECM.
Data quality	Source quality that determines whether the agent's outputs are trustworthy or confidently wrong.
Records governance	Disposition of the agent's own outputs, which are themselves federal records.
Foundation model	The reasoning engine — the most discussed layer and the easiest to source.

FCI Advisory framework, sized to the federal workforce-AI deployment pattern. The model is the easiest of the six layers.

Three of the six layers — integration, content management, and records governance — are where the hardest federal-specific engineering lives, and where most attention does not. A program that scopes and funds these layers alongside model selection ships; a program that treats them as downstream concerns stalls. This is the same pattern FCI documents in its companion report on the federal integration layer: the binding constraint is below the model, not at it.

> *A grievance filed against an agent's decision creates a question no commercial AI deployment has had to answer at scale: who is responsible for the decision, the model or the operator? The answers being negotiated now will set federal precedent.*

04 What program leaders should fund

The implication for a federal workforce-AI program leader is concrete: fund the layers where the program actually succeeds or fails, and evaluate vendors on those layers

rather than on the model benchmark. The table below reframes a conventional AI program budget around where the risk lives.

WHAT TO FUND, AND WHY

Fund this	Because
Integration depth	The middleware that lets the agent reach federal systems of record is the binding constraint on whether the program ships.
Content governance	The records environment the agent operates against determines retrieval quality and records compliance.
Data-quality remediation	Source quality determines whether the agent hallucinates; this is a prerequisite, not a downstream concern.
Union engagement & change management	These programs deploy in unionized, mission-critical environments; the transformation work decides delivery as much as the technology.
Program governance / PMO	Multi-year, multi-vendor, multi-stakeholder programs need deliverables management as a first-class function.

FCI Advisory framework, derived from federal workforce-AI program observation.[4]

What this rules in and out

Four strategic conditions reshape how federal technology leadership should think about AI program design through 2027 and into the next budget cycle:

- **The integration and content layer is where federal workforce AI succeeds or fails.** Foundation-model performance is becoming a baseline; the differentiators are integration depth, content-management governance, and data-quality remediation. Programs that scope and fund these alongside model selection ship; programs that treat them as downstream concerns stall.

- **Vendor evaluations led by model benchmarks are evaluating the wrong layer.** Replace model-benchmark scoring with integration-depth scoring: federal-specific middleware connectors, FedRAMP boundary handling, Documentum and content-integration patterns, audit-trail completeness, records-disposition automation.
- **The deliverables-management challenge is the hidden constraint.** These programs are multi-year, multi-vendor, multi-stakeholder, deployed in unionized mission-critical environments. Program governance, union engagement, change management, and integration sequencing determine whether the program produces what it was supposed to.
- **The pattern will propagate faster than agencies expect.** The labor-relations precedents, federal-specific training conventions, and records-governance patterns being set now will be inherited by federal civilian agencies, defense workforce systems, and commercial organizations with unionized workforces. Watching the sectors furthest down the curve is the cheapest competitive intelligence available.

> *The decision is not whether to deploy agentic AI in workforce systems — that decision has been made and the procurements are in market. It is whether the integration and content layer underneath is being scoped, funded, and governed as a first-class engineering concern, or whether the program will discover the gap at deployment time and pay for it twice.*[4]

Put this thinking to work.

FCI Advisory helps federal agencies build the integration, content, and governance layers underneath workforce AI — the layers that decide whether deployments ship. McLean, Virginia.

Endnotes & sources

1. Federal AI use-case inventories are published annually by agencies under the Advancing American AI Act and OMB direction. The cross-agency aggregate at ai.gov/ai-use-cases is the public reference point.
2. OMB Memorandum M-24-10, "Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence," March 28, 2024. M-24-10 distinguishes "rights-and-safety" AI (higher compliance burden) from administrative AI (lower burden). Most workforce/HR applications fall in the administrative category.

3. NIST AI Risk Management Framework (AI RMF 1.0), National Institute of Standards and Technology, January 2023.
4. Aggregate observations cited in this brief are drawn from FCI Advisory's engagement base across federal workforce-AI and integration programs, FY24-Q4 through FY26-Q1. Quarterly and use-case shares are illustrative of the directional pattern, not point-precise extracts.