DECISION STEWARDSHIP SERIES | FEDERAL AI COMPLIANCE

# Decision Stewardship:

## Why Federal AI Compliance Demands Mission-Specific Decision Support, Not Generic Copilots

*Arthur Billingsley* *February 2026*

## Executive Summary

Federal agencies face a critical inflection point. OMB Memoranda M-25-21 and M-25-22, issued in April 2025 under Executive Order 14179, impose concrete governance, risk management, and procurement requirements on every agency deploying artificial intelligence. Chief AI Officers must establish governance boards, publish formal AI policies, maintain public use case inventories, and implement minimum risk management practices for high-impact systems. These deadlines span from late 2025 through 2026, and they are not optional.

Yet the market response has been overwhelmingly generic. Major vendors are pitching general-purpose AI copilots (chatbots, summarizers, broad-spectrum assistants) as though compliance is a feature toggle rather than an architectural commitment. The result is a widening gap between what agencies are buying and what they are required to document, govern, and defend.

This article introduces the concept of ***decision stewardship***: the discipline of designing, deploying, and governing AI systems that produce auditable, bounded, explainable decision support rather than open-ended generative outputs. Decision stewardship is not a product. It is an operational framework that aligns AI capability with the documentation, risk management, and accountability requirements that OMB now mandates.

> **Key Takeaway**
> Generic AI copilots generate outputs. Decision stewardship produces decision packages: auditable artifacts that include model provenance, data rights documentation, risk controls, constraint boundaries, and human oversight records. Agencies that deploy AI without this distinction will face compliance gaps they cannot close retroactively.

## 1. The Compliance Landscape: What OMB Actually Requires

The dual issuance of M-25-21 (“Removing Barriers to American Leadership in Artificial Intelligence”) and M-25-22 (“Driving Efficient Acquisition of Artificial Intelligence in Government”) reflects a deliberately paired strategy: accelerate adoption while imposing governance

© 2026 COGNOSCERE LLC. All rights reserved. 

guardrails. Understanding both memos together reveals requirements that most vendor pitches quietly omit.

## 1.1 M-25-21: Governance, Risk Management, and Accountability

M-25-21 establishes three interlocking mandates:

- **AI Governance Boards:** Every agency must stand up a governance board with the Chief AI Officer as a principal participant. These boards are not advisory; they hold decision authority over AI use case approval, risk acceptance, and policy enforcement.
- **AI Use Case Inventories:** Agencies must maintain and publicly report inventories of all AI use cases. This is not a checkbox exercise. Each entry must describe the system's purpose, data inputs, decision scope, risk classification, and governance status.
- **Minimum Risk Management Practices:** High-impact AI systems (those affecting individual rights, safety, or access to government services) must meet minimum risk management practices that align with the NIST AI Risk Management Framework. This includes impact assessments, ongoing monitoring, bias testing, and documented human oversight mechanisms.

Critically, M-25-21 requires agencies to **delegate risk acceptance to appropriate officials** for swift action. That requirement deserves close attention: someone must sign their name to a risk determination, which requires documentation sufficient to justify that signature. A general-purpose chatbot produces no such documentation. The official who signs a risk acceptance for an undocumented AI system is accepting personal accountability with no evidentiary foundation.

## 1.2 M-25-22: Procurement as a Governance Mechanism

M-25-22 converts governance requirements into acquisition requirements. Among its mandates:

- **Data Protection:** Vendors must safeguard government data and are explicitly prohibited from using non-public agency data to train commercially available AI models.
- **Vendor Lock-In Prevention:** Contracts must avoid proprietary lock-in, ensuring agencies can migrate between providers or bring capabilities in-house.
- **Transparency and Disclosure:** Vendors must disclose unanticipated AI use. If a subcontractor deploys an AI system the prime did not document, the contract is potentially in breach.
- **Cross-Functional Acquisition Teams:** AI procurement must involve technical, legal, privacy, and mission stakeholders, not just contracting officers and vendor sales teams.

> **Procurement Reality Check**
> M-25-22 effectively means that any AI system acquired for federal use must be accompanied by

© 2026 COGNOSCERE LLC. All rights reserved. 

documentation that supports the governance requirements of M-25-21. If the vendor cannot provide model cards, data rights documentation, risk assessments, and audit-ready artifacts at the point of delivery, the agency has a compliance gap from day one.

## 2. The Generic Copilot Problem

The current vendor landscape is dominated by a seductive pitch: deploy a general-purpose AI assistant, integrate it with your data, and let users ask questions. The implicit promise is that capability equals compliance, that a powerful enough model wrapped in an enterprise interface satisfies the governance requirements.

It does not. Here is why.

### 2.1 Outputs vs. Decision Packages

A generic copilot produces *outputs*: summaries, drafted emails, data visualizations, code suggestions. These outputs have no inherent provenance chain. They do not document which data was used, what constraints were applied, what risk boundaries were enforced, or whether a human reviewed the result before action was taken.

A decision support system operating under a stewardship framework produces *decision packages*: bounded, documented artifacts that include not just the recommendation but the entire context required to evaluate, audit, and defend that recommendation.

| Dimension | Generic Copilot | Decision Stewardship System |
|---|---|---|
| Output Type | Open-ended text generation | Bounded decision recommendation with rationale |
| Data Provenance | Undocumented or opaque | Logged: sources, versions, access rights, classification |
| Model Documentation | Vendor-provided (if any) | Model card, version, validation results, drift monitoring |
| Constraint Boundaries | User-defined (or absent) | Pre-configured mission-specific guardrails |
| Risk Controls | Generic content filters | Decision-type-specific risk thresholds and escalation rules |
| Audit Trail | Chat logs (if retained) | Structured audit record: inputs, model, constraints, output, reviewer |
| Human Oversight | Optional review | Mandatory review gates with documented sign-off |

© 2026 COGNOSCERE LLC. All rights reserved. 

| | | |
|---|---|---|
| Compliance Posture | Retroactive documentation burden | Compliance-by-design from deployment |

### 2.2 The NIST Alignment Gap

The NIST AI Risk Management Framework (AI RMF 1.0, January 2023, with ongoing updates) provides the de facto standard for federal AI governance. Its four core functions (Govern, Map, Measure, Manage) describe an iterative cycle of risk identification and mitigation.

Generic copilots map poorly to this framework:

- **Govern:** Copilots lack built-in governance structures. Policies must be retrofitted through usage guidelines, which are routinely ignored or circumvented.
- **Map:** The AI RMF requires organizations to contextualize AI risks relative to their specific mission. A general-purpose assistant, by definition, lacks mission-specific context.
- **Measure:** Measuring bias, accuracy, and reliability requires defined decision types with testable outcomes. Open-ended generation produces outputs that resist systematic measurement.
- **Manage:** Risk mitigation requires bounded operational parameters. A copilot that can answer any question about any topic cannot enforce the kind of decision-specific constraints that NIST envisions.

The March 2025 NIST update further expanded threat categories to include poisoning attacks, evasion attacks, data extraction, and model manipulation. These threats are significantly harder to detect and mitigate in systems without bounded operational envelopes.

The gap between what copilots provide and what NIST requires is not a tuning problem. It is an architectural problem. Closing it requires a fundamentally different design philosophy.

## 3. What Decision Stewardship Looks Like in Practice

Decision stewardship is not a single technology. It is a design philosophy and operational discipline. Here is what it looks like when implemented.

### 3.1 The Decision Package

Every AI-assisted decision under a stewardship framework generates a structured decision package containing:

- **Decision Request:** What question was asked, by whom, under what authority, and within what mission context.

© 2026 COGNOSCERE LLC. All rights reserved. 

- **Data Inputs:** Which datasets were consulted, their classification levels, currency, provenance, and any access restrictions.
- **Model Identity:** Which model(s) were used, their version numbers, validation status, known limitations, and most recent test results.
- **Constraint Set:** What guardrails were active, including topic boundaries, confidence thresholds, escalation triggers, and prohibited output categories.
- **Raw Output:** The unmodified model output before any post-processing or human review.
- **Human Review Record:** Who reviewed the output, what modifications were made, what was the final decision, and what was the documented rationale.
- **Risk Assessment:** The system's self-assessed confidence level, flagged uncertainties, and any triggered risk escalations.
- **Audit Metadata:** Timestamps, system identifiers, session hashes, and retention classification.

> **Why This Matters for DAWIA-Trained Acquisition Professionals**
>
> If you have worked in defense acquisition, this structure should look familiar. It mirrors the documentation rigor of a System Engineering Technical Review or a decision brief for a Milestone B review. Decision stewardship applies the same discipline to AI-assisted decisions that the acquisition community already applies to weapons system decisions. The principle is identical: consequential decisions require documented, reviewable, defensible rationale.

## 3.2 Bounded Operational Envelopes

A decision stewardship system does not answer arbitrary questions. It operates within a **defined operational envelope**: a bounded set of decision types, data sources, and output parameters that align with a specific mission function.

Consider an AI system supporting benefits adjudication. Under a stewardship model, that system would be constrained to the statutory criteria for the specific benefit program, the applicant's submitted documentation and verified government records, recommendations within defined decision categories (approve, deny, refer for additional review), and mandatory escalation for any case flagging protected-class disparate impact indicators.

This is the opposite of a copilot. A copilot's value proposition is breadth: it can do anything. A stewardship system's value proposition is *precision*. It does exactly what the mission requires, documents everything it does, and refuses to operate outside its authorized boundaries.

## 3.3 Continuous Compliance Architecture

© 2026 COGNOSCERE LLC. All rights reserved. 

Under the decision stewardship model, compliance is not an annual audit exercise. It is a continuous architecture that includes:

- **Real-Time Audit Logging:** Every interaction generates an auditable record in a format that feeds directly into the agency's AI Use Case Inventory and governance reporting.
- **Automated Drift Detection:** Model performance is continuously monitored against validated baselines, with automatic alerts when outputs begin diverging from expected parameters.
- **Bias Monitoring Dashboards:** Decision patterns are tracked across demographic categories with statistical testing for disparate impact, not as a periodic review, but as a live operational metric.
- **Version-Controlled Model Updates:** No model change is deployed without documented testing, risk assessment, and governance board notification, mirroring the configuration management discipline of any mission-critical system.

# 4. The Acquisition Case for Decision Stewardship

For program managers and contracting officers operating under M-25-22's procurement requirements, decision stewardship changes the acquisition conversation in three fundamental ways.

## 4.1 Evaluation Criteria That Work

Traditional AI procurement evaluations focus on model capability: accuracy benchmarks, training data size, inference speed. Under M-25-22, evaluation criteria must also address:

- **Documentation Completeness:** Can the vendor deliver model cards, data rights documentation, and risk assessments at contract delivery?
- **Governance Integration:** Does the system produce outputs compatible with the agency's AI Use Case Inventory reporting requirements?
- **Audit Architecture:** Is the audit trail a first-class architectural element, or a logging afterthought?
- **Portability:** If the agency changes vendors, do the governance artifacts transfer, or are they locked in the vendor's proprietary format?

## 4.2 Total Cost of Compliance

Generic copilots appear cheaper at acquisition but impose significant downstream compliance costs. The agency must separately build governance documentation for every use case, audit infrastructure that the copilot does not provide, risk management processes external to the

© 2026 COGNOSCERE LLC. All rights reserved. 

system, bias testing and monitoring capabilities, and human oversight workflows with their associated documentation.

A stewardship-designed system includes these capabilities as integral components. The total cost of compliance (not just the license fee) consistently favors purpose-built decision support over retrofitted general-purpose assistants.

### 4.3 Risk Transfer vs. Risk Acceptance

When an agency deploys a generic copilot, it accepts the entirety of the governance risk. The vendor provides a tool; the agency must figure out compliance. When an agency acquires a decision stewardship system, a significant portion of the compliance burden is transferred to the architecture itself. The system's design enforces constraints, generates documentation, and maintains audit trails by default.

This is not a theoretical distinction. When the agency's designated official must sign a risk acceptance determination under M-25-21, the question is simple: would you rather sign based on a chat log from a general assistant, or based on a structured decision package with full provenance documentation?

## 5. From Concept to Implementation: A Practical Roadmap

Transitioning from generic AI deployment to decision stewardship does not require discarding existing investments. It requires a structured approach to layering governance architecture onto AI capabilities.

### Phase 1: Decision Inventory and Classification

Before deploying any AI system, catalog the decisions the system will support. Classify each by impact level (routine, significant, high-impact), data sensitivity, and statutory or regulatory constraints. This inventory becomes the foundation for both the operational envelope design and the AI Use Case Inventory required by M-25-21.

### Phase 2: Governance Architecture Design

For each decision class, define the constraint set, required data sources, acceptable model types, risk thresholds, escalation rules, and human oversight requirements. This is the engineering work that separates stewardship from improvisation.

### Phase 3: Decision Package Implementation

Implement the technical infrastructure to generate, store, and retrieve decision packages. This includes audit logging, model version tracking, data provenance recording, and human review workflow management.

© 2026 COGNOSCERE LLC. All rights reserved. 

### Phase 4: Continuous Monitoring and Governance Reporting

Deploy bias monitoring, drift detection, and performance dashboards that feed directly into governance reporting. Establish feedback loops between operational data and governance board reviews.

> **Implementation Reality**
> None of this is exotic technology. The components (structured logging, model registries, workflow management, statistical monitoring) are mature capabilities. What is missing in most federal AI deployments is the architectural intent to assemble them into a coherent governance framework. Decision stewardship provides that architectural intent.

## 6. Conclusion: The Stewardship Imperative

The federal AI compliance landscape is not going to simplify. OMB's requirements reflect a bipartisan recognition that AI in government requires accountability structures commensurate with its decision-making power. Agencies that deploy generic copilots without governance architecture are building compliance debt that compounds with every interaction.

Decision stewardship offers a different path: AI systems designed from the ground up to produce not just useful outputs, but **defensible decisions**. Systems that document their own provenance, enforce their own boundaries, and generate the governance artifacts that compliance demands.

The distinction between a copilot and a decision stewardship system is the distinction between a tool and an institution. Tools require their users to maintain discipline. Institutions embed discipline into their architecture.

Federal agencies deserve institutions. The time to build them is now.

## References

*Office of Management and Budget.* Memorandum M-25-21: Removing Barriers to American Leadership in Artificial Intelligence. April 3, 2025.

*Office of Management and Budget.* Memorandum M-25-22: Driving Efficient Acquisition of Artificial Intelligence in Government. April 3, 2025.

*Executive Office of the President.* Executive Order 14179: Removing Barriers to American Leadership in Artificial Intelligence. January 2025.

© 2026 COGNOSCERE LLC. All rights reserved. 

*National Institute of Standards and Technology.* AI Risk Management Framework (AI RMF 1.0). NIST AI 100-1. January 2023.

*National Institute of Standards and Technology.* Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. NIST AI 600-1. July 2024.

*National Institute of Standards and Technology.* Cybersecurity Framework Profile for AI Risk Management. NISTIR 8596. Preliminary Draft, December 2025.

*Defense Acquisition University.* Defense Acquisition Workforce Improvement Act (DAWIA) Certification Standards. Current edition.

## About the Author

Arthur Billingsley is a retired Navy Commander and founder of COGNOSCERE LLC, an IT management consulting firm specializing in AI compliance, decision support architecture, and digital transformation for federal and regulated industries. He has more than a decade's experience as adjunct and assistant professor in computer and electrical engineering. With deep expertise in both defense acquisition (DAWIA Level II) and commercial technology deployment, he bridges the gap between regulatory requirements and operational AI governance.

**Contact:** info@cognoscere.com | **Series:** This article is part of COGNOSCERE's Decision Stewardship Series on AI governance for federal and regulated industries.

© 2026 COGNOSCERE LLC. All rights reserved. 