FERZ INC.

AI GOVERNANCE EXECUTIVE GUIDE SERIES · VOLUME 4

AI Governance for Government Contractors

Winning Work, Protecting Relationships, Surviving Audits

A practical guide for delivering AI services under OMB M-25-21 and M-25-22, with contractor evidence packages, subcontractor governance, and Inspector General audit readiness.

> *Governance is enforced at the authorization boundary: the point where an output becomes an effect-bearing action, and where evidence, not explanations, must exist.*

PUBLISHER
FERZ, Inc.

VERSION
1.1 · June 2026

FOR

Chief Technology Officers · Program Managers · Capture Managers · Proposal Teams · Compliance Officers · AI and ML Technical Leads

FRONT MATTER

License and Terms of Use

Copyright and License

Attribution

FERZ, Inc., "AI Governance Executive Guide: Government Contractors," 2026. Available at ferz.ai.

Public Reference Framework

FERZ, Inc. publishes this as a public, vendor-neutral reference framework based entirely on public sources. It does not use nonpublic government information, does not speak for or on behalf of any federal agency, and does not imply government endorsement.

Framework Scope and Citation Currency

This guide is a foundational governance framework, not an exhaustive policy tracker or current-law reference. It cites statutes, regulations, memoranda, standards, and guidance only where they materially affect the general governance model, acquisition posture, evidence requirements, oversight structure, or audit-readiness framework. Use-case-specific, agency-specific, jurisdiction-specific, or rapidly changing requirements may apply and should be verified separately before reliance or implementation.

Not Legal, Procurement, or Compliance Advice

This document is provided as a general reference. The clause language, checklists, proposal language, and procedures in this guide are illustrative and are not legal, procurement, or compliance advice. Contractors should adapt any language to their own contracts, authorities, and counsel review. This document is provided "as is," without warranties of any kind.

> **Commercial Scope Notice**
>
> Commercial services, tooling, and implementation support derived from this framework are offered separately by FERZ, Inc. or its authorized partners. This notice does not restrict rights granted under the license.

Limitation of Liability

To the maximum extent permitted by law, FERZ, Inc. is not liable for any direct, indirect, incidental, consequential, special, exemplary, or other damages arising from use of, reliance on, adaptation of, or implementation based on this publication. Users are responsible for obtaining appropriate legal,

regulatory, procurement, ethics, technical, and compliance review before applying any framework, checklist, clause language, or procedure described in this guide.

EXECUTIVE AT-A-GLANCE

What This Guide Does

Audience

Chief Technology Officers, Program Managers, Capture Managers, Proposal Teams, Compliance Officers, and AI and ML technical leads at federal contractors, defense primes, systems integrators, and government IT service providers.

Core Thesis

> **Governance is enforced at the authorization boundary: the point where an output becomes an effect-bearing action, and where evidence, not explanations, must exist.**

Contractor Readiness Thesis

Contractors best positioned to compete are those that can produce authorization-ready evidence, not merely promise governance.

The Shift

Federal buyers are getting more specific about AI governance. They are adding AI clauses to solicitations, requiring evidence packages, reserving verification and reclassification rights, and reserving the authority to suspend AI operations pending investigation. Contractors that are ready for this shift can respond on the contract's schedule. Contractors that are not will reconstruct governance after award, under time pressure.

Five Readiness Payoffs

PAYOFF	HOW IT WORKS
Responsive to AI clauses	When a solicitation requires AI governance, a ready contractor already meets the requirement and can respond on schedule.
Reusable evidence	Pre-built evidence packages are reusable across engagements, which lowers proposal and delivery effort.
Audit durability	Maintained documentation supports timely and complete responses to Inspector General and government requests.
Relationship protection	Disciplined governance reduces the AI incidents that surface through complaints, FOIA, or oversight.
Defensible positioning	Demonstrated readiness supports a credible value position rather than a promise to build the capability later.

What Is Appearing in Federal Contracts

Government buyers are increasingly including AI-specific requirements:

- **AI System definitions** that reach primary-purpose AI and AI integrated into a business process, while excluding AI that is only incidentally present.
- **The Governed Unit concept**, under which each use case is governed independently rather than as one model.
- **Evidence packages** covering training data provenance, validation results, bias testing, and incident history, scaled to risk tier.
- **Verification rights**, including pre-award testing and post-award monitoring and audit access.
- **Suspension authority**, under which the government can stop an AI operation pending investigation.

This guide shows you how to be ready for each of these.

EXECUTIVE SUMMARY

The Era of Prove It

The federal AI governance landscape has shifted. Agencies are implementing OMB M-25-21 and M-25-22, designating Chief AI Officers, and flowing requirements down to contractors. The posture of "trust us, our AI works" is giving way to "show the evidence."

This is not only a compliance burden. It is a position you can prepare for. Many competitors are not ready: they do not maintain evidence packages, they cannot produce training data documentation on request, and they have not adopted the Governed Unit concept. When the new contract requirements appear, that gap becomes visible.

> REGULATORY BRIDGE
>
> M-25-21 consolidates the prior rights-impacting and safety-impacting categories into the current high-impact AI framework. Existing rights and safety concepts remain relevant, but the operative category for this guide is high-impact AI.

The Government Buyer's New Playbook

Federal program managers now have guidance on what to require from contractors. They are being directed to:

- Define AI by capability and function, reaching systems established primarily for AI and AI integrated into another business process.
- Require disclosure of AI used in performance, including AI use that was not anticipated at award.
- Require evidence packages scaled to risk tier.
- Reserve the right to reclassify an AI use case to a higher tier.
- Conduct verification activities, including pre-award testing and ongoing monitoring, rather than accept assertions.
- Reserve the authority to suspend AI operations pending investigation.

This is not theoretical. M-25-22 requirements apply to contracts awarded under solicitations issued on or after September 30, 2025, and to options to renew or extend exercised on or after October 1, 2025. Expect AI clauses first in new awards and recompetes, then as modifications on incumbents.

What This Guide Provides

- **Contract clause decoder:** what the AI clauses mean and how to respond.
- **Evidence package templates:** what to pre-build so you are ready at award.
- **Tier classification guidance:** how to self-assess honestly and avoid reclassification surprises.
- **Subcontractor management:** how to obtain what you need from AI vendors.
- **Proposal language:** how to make governance readiness a substantiated discriminator.
- **Audit readiness:** how to respond completely and on time.

The Bottom Line

Government buyers are learning to ask a single question: how do you know your AI is governed? The contractors who can answer with evidence, not assertions, are the ones positioned to win and to keep the work.

PART I

The New Landscape

CHAPTER 1

What Government Buyers Now Expect

Federal acquisition of AI is maturing. Federal program managers, Contracting Officers, and Contracting Officer's Representatives are increasingly being asked to account for AI oversight. M-25-21 and M-25-22 are being implemented. Inspector General offices are adding AI to their audit plans. This chapter explains what is changing and why it matters to your business.

The M-25-21 and M-25-22 Effect

OMB M-25-21 requires agencies to govern AI, including AI that contractors operate, and M-25-22 governs how agencies acquire it. The implications for you:

- **Agency AI inventories include your systems.** If you operate governed AI to perform contract requirements, it belongs in the agency's accounting of AI in use.
- **High-impact classification flows down.** If your AI is high-impact under M-25-21, heightened minimum practices apply, and they reach the contractor operating the system.
- **Minimum practices become contract requirements.** Agencies must ensure contractor AI meets the standard, and the mechanism is the contract.

> AUTHORIZATION BOUNDARY
>
> OMB M-25-21 requires agencies to apply minimum practices to high-impact AI, including pre-deployment testing, impact assessment, ongoing monitoring, human oversight, intervention, accountability, and appropriate fail-safes. For the contractor operating that AI, the authorization boundary is where those requirements become enforceable: a high-impact AI action releases only on an ALLOW verdict, and if the authorization artifact cannot be produced, the action fails closed.

Agency-specific requirements. M-25-21 and M-25-22 are the floor, not the ceiling. Individual agencies may impose stricter controls through agency-specific AI policies, directives, acquisition procedures, or contract terms. Know your customer's specific requirements and verify current agency supplements before relying on any general framework.

The New Contract Clauses

Government buyers are adding AI-specific clauses to the Performance Work Statement or Statement of Work. Here is what to expect.

AI System Definition

For this guide, an AI system includes an AI capability acquired, developed, integrated, or operated to perform contract requirements, including primary-purpose AI systems, AI services, embedded AI used in a governed workflow, and subcontractor-provided AI. Common commercial products with incidental embedded AI are not treated as governed AI merely because they contain AI features, unless the AI capability is directed, required, or materially relied upon to perform the contract or support an effect-bearing function.

This scope is broad enough to reach the AI that matters for performance, and bounded enough that it does not sweep in every commercial product with an AI feature. The right move is to inventory honestly against that line, not to argue that effect-bearing AI is something other than AI.

Governed Unit Concept

A Governed Unit is an AI system performing a specific use case in a specific workflow. The same model in two different contexts is two Governed Units, each independently classified and governed. You cannot validate one model and treat the question as closed. Each deployment context carries its own governance.

Evidence Package Requirements

For Tier 1, high-impact AI, expect requirements to provide:

- Training data documentation, including sources, demographics, selection criteria, and limitations.
- Validation results, including demographic subgroup analysis.
- Civil-rights, safety, and fairness testing methodology and results.
- Human oversight procedures, with evidence that the oversight is meaningful.
- Incident history and corrective actions.
- Ongoing monitoring procedures and data.

Government Verification Rights

The government is no longer only accepting assertions. New clauses provide for:

- Pre-award testing of the system.
- Post-award monitoring and the right to evaluate performance, risk, and effectiveness.
- Access to documentation and to personnel for interviews.
- Access to test environments for independent validation.
- A defined response window for documentation requests, as specified in the contract.
- Authority to suspend AI operations pending investigation.

LLM Transparency Documentation

OMB M-26-04 adds a current documentation layer for procured LLMs. Contractors should expect agencies to request acceptable-use policies, model, system, or data cards where available, end-user resources, feedback mechanisms, evaluation results, and documentation of controls relevant to the procured use. Treat these materials as part of the contractor evidence package, not as marketing collateral.

READINESS PAYOFF

Contractors that maintain current evidence packages can respond to documentation requests without reconstructing governance after award.

CHAPTER 2

The Competitive Landscape

AI governance readiness is becoming a differentiator. This chapter explains how to position against competitors who are not ready, using evidence rather than claims.

Where Most Contractors Are Today

An honest assessment: many federal contractors are not ready for the new AI governance requirements. Common gaps:

- **No AI inventory.** They do not actually know what AI is running on their contracts.
- **No training data documentation.** They cannot explain where the data came from or its limitations.
- **No subgroup testing.** They validated overall performance but not across demographic groups.
- **Checkbox human oversight.** Someone approves, but does not actually review.
- **No incident history.** Problems occurred but were not documented as AI incidents.
- **Subcontractor black boxes.** They use vendor AI but cannot obtain governance documentation.

What Governance Ready Looks Like

CAPABILITY	NOT READY	READY
AI inventory	"We will figure out what we have"	Complete inventory with tier classifications
Training data	"The vendor handles that"	Documented provenance, demographics, limitations
Validation	"It passed our tests"	Subgroup analysis, civil-rights, safety, and fairness testing, limitations documented
Human oversight	"Someone approves it"	Defined authority, training, time allocation, accountability
Incidents	"We fix problems"	Documented incident history with root cause and remediation
Subcontractors	"They are compliant"	Evidence packages on file, audit rights in subcontracts

Making Governance a Proposal Discriminator

When responding to solicitations with AI governance requirements, substantiate readiness rather than promise it.

Proposal Language That Substantiates

- "Our AI governance framework is aligned with OMB M-25-21 and NIST AI RMF."
- "We maintain evidence packages for all Tier 1 AI systems, including the artifacts listed in our governance appendix."
- "Our standard subcontracts include AI governance flow-down and audit rights."
- "We can provide the required evidence package within the contract's specified response window."
- "Our AI incident response procedure has been tested and documented."

Proof Points to Include

- A redacted sample evidence package from a comparable engagement.
- An AI governance organizational chart with named roles.
- A list of pre-validated AI tools with governance documentation.
- Training records showing AI governance competency.

> **READINESS PAYOFF**
>
> A claim backed by an attachable artifact is verifiable. A claim that the capability will be developed after award is not, and a sophisticated evaluator reads the difference.

PART II

Building Your Governance Capability

CHAPTER 3

Evidence Packages

Evidence packages are the core deliverable of AI governance. This chapter covers what to build, how to maintain it, and how to make it reusable across contracts.

Tier 1 Evidence Package Contents

For Tier 1 high-impact AI, especially systems whose outputs are a principal basis for effect-bearing rights, safety, mission, or infrastructure decisions, build and maintain:

ARTIFACT	CONTENTS	UPDATE FREQUENCY
System Description	Architecture, intended use, decision authority, limitations, contraindications	On change
Training Data Doc	Sources, demographics, selection criteria, known gaps, provenance chain	On retraining
Validation Report	Performance metrics overall and by subgroup, test methodology, limitations	Annual and on change
Civil Rights, Safety, and Fairness Testing	Methodology, performance results across relevant populations or operational contexts, adverse-impact analysis where applicable, and remediation steps	Annual and on change
Human Oversight SOP	Reviewer roles, authority, training requirements, time allocation, escalation	Annual review
Incident Log	All incidents, root cause, corrective actions, government notifications	Continuous
Monitoring Dashboard	Metrics tracked, thresholds, drift detection, alert procedures	Continuous
Authorization Artifact Evidence	For effect-bearing actions, sample or production authorization artifacts showing the pre-release verdict, the applicable policy state, the authority chain, the provenance condition, and the replay basis	Continuous
LLM Transparency Documentation, where applicable	Acceptable-use policies, model/system/data cards where available, evaluation methodology, feedback mechanisms, and documentation of controls relevant to the procured use	On change

Building Evidence Packages Efficiently

Evidence packages should not be one-off efforts. Build infrastructure:

- **Templates:** standardized templates for each artifact type, pre-populated where possible.
- **Automation:** automated collection of monitoring data, incident logs, and validation metrics.
- **Version control:** all artifacts versioned with clear change history.
- **Central repository:** a single source of truth accessible to program teams and to auditors.
- **Review calendar:** scheduled reviews to keep documents current.

What Government Reviewers Look For

When evaluating your evidence package, government reviewers and IG auditors look for:

- **Completeness:** all required artifacts present, with no obvious gaps.
- **Currency:** documents reflect the current system state, not outdated versions.
- **Specificity:** concrete detail, not generic boilerplate.
- **Honesty about limitations:** documented constraints are more credible than claims of perfection.
- **Traceability:** a clear path from requirements to testing to deployment.

Red Flags That Trigger Deeper Review

- Generic descriptions that could apply to any AI system.
- Validation results that are too clean, with no limitations documented.
- Human oversight procedures with no time allocation or training requirements.
- Empty incident logs, since "no incidents" is less credible than "incidents with resolution."
- Subcontractor AI with no governance documentation.

> READINESS PAYOFF
>
> An evidence package that honestly documents limitations is more credible than one that claims perfection. Reviewers know AI has limitations; they want to see that you have identified and managed them.

CHAPTER 4

Tier Classification

Classifying your AI correctly is critical. Under-classify and the government will reclassify, with delay and credibility cost. Over-classify and you carry unnecessary overhead. Get it right the first time.

The Four Tiers

Tier 1 corresponds to high-impact AI as defined by OMB M-25-21. Tiers 2 to 4 are an oversight-calibration scheme for AI that falls below the high-impact threshold.

TIER	CRITERIA	EXAMPLES	WHAT YOU MUST PROVIDE
Tier 1	High-impact: principal basis for effect-bearing decisions or actions	Eligibility, threat detection, autonomous systems	Full evidence package; government approval before deployment
Tier 2	Operational impact, federal employee facing	Caseworker tools, document processing	Validation summary, annual review, monitoring
Tier 3	Internal, no citizen impact	Meeting transcription, internal search	Basic documentation, security-baseline alignment
Tier 4	Personal productivity, no government data	General research tools	Acceptable use policy only

Automatic Tier 1 Triggers

These conditions make any AI system Tier 1, regardless of your assessment:

- Outputs substantially factor into decisions affecting civil rights or equal opportunity.
- Outputs substantially factor into decisions affecting human safety or critical infrastructure.
- The system produces citizen-facing determinations, even with human review.
- The system involves classified information, controlled unclassified information, or sensitive personally identifiable information in decision logic.
- The system has autonomous or semi-autonomous action capability.

Do not argue around these. If a trigger applies, classify as Tier 1.

The Governed Unit Trap

Governance applies to each Governed Unit, the AI plus use case plus workflow, not to the model in the abstract. Common mistakes:

- **Wrong:** "Our NLP model is Tier 3 because it is just text processing."
- **Right:** "Our NLP model processing benefits applications is Tier 1; the same model doing internal search is Tier 3."

The government can reclassify any Governed Unit to a higher tier. If you have under-classified, you will have to produce Tier 1 documentation on short notice, or face suspension.

> READINESS PAYOFF
>
> Classify honestly. A Tier 1 classification with solid documentation is stronger than a Tier 3 classification that gets overruled, which carries both rework and credibility cost.

CHAPTER 5

Managing Subcontractors and Vendors

If you use AI from subcontractors or commercial vendors, you are responsible for its governance. "The vendor handles that" is not an acceptable answer to a government auditor.

The Prime Contractor's Problem

As prime, you face the government. When the Contracting Officer asks for an evidence package, you must provide it, even when the AI comes from a subcontractor. When the IG audits, they audit you. Common failure modes:

- **Vendor black box:** a commercial AI vendor will not provide training data documentation.
- **Subcontractor resistance:** a sub claims governance artifacts are proprietary.
- **Missing flow-down:** your subcontract does not include AI governance requirements.
- **No audit rights:** you cannot actually verify the sub's claims.

What to Require from Subcontractors

Include in every subcontract involving AI:

REQUIREMENT	TIER 1	TIER 2	TIER 3
AI system description and intended use	Required	Required	Required
Training data provenance documentation	Required	Required	Not applicable
Validation results with subgroup analysis	Required	Summary	Not applicable
Civil-rights, safety, and fairness testing results	Required	Risk-based	Not applicable
Incident history and remediation	Required	Required	Requested
Authorization artifact evidence for effect-bearing AI actions	Required	Risk-based	Not applicable
Audit and inspection cooperation	Required	Required	Required
Government access rights, flow-through	Required	Required	Required

Flow-down controls should require the subcontractor to evidence that each effect-bearing AI action was authorized before release. Monitoring logs alone are not accepted as evidence that an action was permitted.

Negotiating with Resistant Vendors

Some vendors will push back on governance requirements. Measured responses:

- **"That is proprietary."** Reply: we need governance documentation, not source code. If you cannot provide it, we will find a vendor who can.

- **"No one else asks for this."** Reply: federal contracts now require it. We need vendors who can meet federal requirements.
- **"We are SOC 2 certified."** Reply: SOC 2 does not, by itself, satisfy AI governance evidence requirements. We need AI-specific documentation for this governed use case.
- **"Our AI is too complex to document."** Reply: then it is too complex to deploy on a federal contract.

Building a Governance-Ready Vendor List

- Send AI governance requirements to potential vendors before you need them.
- Evaluate governance capability as part of vendor selection.
- Maintain a list of governance-ready vendors for rapid proposal response.
- Include governance requirements in master agreements, not only in task orders.

> READINESS PAYOFF
>
> A pre-qualified list of governance-ready vendors lets you respond to an AI clause immediately, rather than beginning vendor negotiations after the solicitation drops.

PART III

Operations and Compliance

CHAPTER 6

Day-to-Day Governance Operations

Governance is not a one-time deliverable. It is ongoing operations. This chapter covers the operational rhythm that keeps you compliant and audit-ready.

The Governance Calendar

FREQUENCY	ACTIVITY	OWNER
Continuous	Monitoring data collection, incident logging	Technical team
Monthly	Monitoring review, threshold check, drift assessment	AI governance lead
Quarterly	Evidence package currency review, subcontractor check-in	Program manager
Annual	Full revalidation for Tier 1, bias retesting, policy review	AI governance lead
On change	Impact assessment, re-tiering evaluation, documentation update	Technical team and governance

Incident Management

AI incidents will happen. How you handle them determines whether they become audit findings.

What Counts as an Incident

- AI produces incorrect output affecting a citizen or mission outcome.
- Performance degrades below validated thresholds.
- Bias or disparate impact is identified.
- Human oversight is bypassed, intentionally or accidentally.
- A data quality issue affects AI operation.
- A security event involves the AI system.

Incident Response Requirements

- **Notification:** government contracts typically require notification within a defined window for Tier 1 AI incidents. Know your contract's specific requirement and meet it.
- **Root cause analysis:** determine why it happened, not only what happened.
- **Corrective action:** document what you are doing to prevent recurrence.
- **Evidence preservation:** preserve logs, data, and system state for potential audit.

The Incident Log

Maintain a continuous incident log containing the date and time of the incident and its detection, a description of what happened, an impact assessment, root cause when determined, corrective actions taken, government notification with date and recipient where required, and resolution and closure.

> **READINESS PAYOFF**
>
> A well-documented incident history shows mature governance. An empty incident log raises the question of whether problems are being detected, not whether they exist.

CHAPTER 7

Surviving Audits

Inspector General audits of AI are increasing. This chapter prepares you to respond completely and on time, and to benefit from the scrutiny.

What Auditors Ask For

Based on emerging IG audit practice, expect requests across four areas.

Inventory and Classification

- A complete list of AI systems on this contract.
- Tier classification for each, with rationale.
- Evidence of government disclosure within the required timeframes.

Evidence Packages

- Training data documentation, validation results, and bias testing methodology and results.
- Human oversight procedures.

Operational Records

- Monitoring data for the past twelve months.
- The incident log with all entries, government notifications made, and corrective actions taken.

Subcontractor Oversight

- A list of subcontractors providing AI, evidence packages from each, and your verification activities.

The Documentation Response Window

New contract clauses commonly require documentation within a short window, often a few business days, of a government request. If you cannot respond in time, the consequences include credibility cost with the program office, potential escalation to the Contracting Officer, an audit finding for inadequate documentation, and a contract performance issue. If you have built and maintained your evidence packages, the window is easy to meet. If you have not, it is not.

Common Audit Findings to Avoid

FINDING	HOW TO AVOID IT
Incomplete AI inventory	Know what AI is running. Include embedded AI and subcontractor AI.
Missing evidence packages	Build and maintain packages before you need them.
Inadequate training data documentation	Document provenance, demographics, and limitations.
No subgroup validation	Test performance across demographic groups, not only overall.
Checkbox human oversight	Document authority, training, and time allocation, and show the oversight is meaningful.
Unreported incidents	Log everything. Report per contract requirements.
Subcontractor gaps	Obtain evidence packages from subs. Do not accept "proprietary" as a shield.

READINESS PAYOFF

The best audit outcome is no findings. The next best is findings already remediated. Maintained governance is what puts you in one of those two positions.

PART IV

Building Your Capability

CHAPTER 8

Organizational Readiness

AI governance requires organizational capability, not only documentation. This chapter covers the people, processes, and authority you need.

Key Roles

ROLE	RESPONSIBILITIES	REQUIRED COMPETENCIES
AI Governance Lead	Policy, standards, evidence package oversight, audit coordination	AI basics, federal compliance, program management
Technical AI Lead	Validation, monitoring, incident investigation, documentation accuracy	ML and AI technical depth, validation methodologies
Program AI POC	Contract-specific governance, government liaison, evidence delivery	Contract knowledge, AI awareness, communication
Subcontract Manager	Vendor and sub AI requirements, evidence collection, flow-down	Procurement, AI governance requirements

Authority Requirements

Governance fails when the function lacks authority. Ensure:

- **Tier classification authority:** the AI governance lead can classify and escalate, not only recommend.
- **Deployment gates:** no Tier 1 AI deploys without governance sign-off.
- **Budget:** validation, testing, and documentation have dedicated budget, not budget borrowed from delivery.
- **Escalation path:** a clear path to stop deployment when governance is inadequate.

Competency Development

Your people need to understand AI governance, not only check boxes:

- **AI basics:** what AI is, how it works, and common failure modes.
- **Federal requirements:** M-25-21, M-25-22, agency-specific requirements, and contract clauses.
- **Validation methods:** how to evaluate bias testing, subgroup analysis, and monitoring data.
- **Evidence package contents:** what belongs in each artifact and what reviewers look for.

Process Infrastructure

- Templates for all evidence package artifacts.
- Checklists for tier classification.
- A workflow for government notifications.
- A calendar for periodic reviews.
- A repository for evidence packages with access controls.
- An incident response playbook.

CHAPTER 9

Implementation Roadmap

If you are starting from scratch, here is how to build governance capability in 90 days.

Phase 1: Foundation, Weeks 1 to 4

Weeks 1 to 2: Inventory and Assessment

- Identify all AI on current federal contracts, including embedded and subcontractor AI.
- Assess the current state of documentation for each.
- Identify gaps against Tier 1 evidence requirements.
- Review current subcontracts for AI governance provisions.

Weeks 3 to 4: Structure and Standards

- Assign AI governance roles.
- Establish tier classification criteria.
- Create evidence package templates.
- Draft standard subcontract AI provisions.

Phase 2: Build, Weeks 5 to 8

Weeks 5 to 6: Evidence Packages

- Prioritize Tier 1 and Tier 2 systems.
- Build evidence packages for the highest-risk systems first.
- Request evidence from subcontractors.
- Identify vendor gaps and negotiate solutions.

Weeks 7 to 8: Operations

- Implement monitoring for Tier 1 systems.
- Establish incident logging.
- Create the governance calendar.
- Test the documentation response capability.

Phase 3: Operationalize, Weeks 9 to 12

Weeks 9 to 10: Integration

- Integrate governance into the proposal process.
- Add AI governance to subcontract templates.
- Train program teams on governance requirements.
- Brief capture teams on substantiated positioning.

Weeks 11 to 12: Validation

- Conduct a mock audit.
- Remediate gaps identified.
- Document lessons learned.
- Establish a continuous improvement process.

> READINESS PAYOFF
>
> A disciplined ninety-day effort can move a contractor from ad hoc documentation toward governance readiness, provided the relevant systems, vendors, and evidence sources are accessible. A contractor that waits will need that same time after the solicitation drops, when there is no longer time to spend.

CHAPTER 10

Maturity Progression

Governance capability develops over time. Here is the progression from basic compliance to independently verifiable governance.

LEVEL	NAME	WHAT IT LOOKS LIKE	BUSINESS VALUE
1	Reactive	Build documentation when asked	Can respond, slowly, to requirements
2	Structured	Standard templates, assigned roles	Faster response, consistent quality
3	Managed	Pre-built packages, proactive monitoring	Proposal discriminator, audit-ready
4	Optimized	Automated collection, predictive risk identification	Strong positioning, minimal overhead
5	Independently Verifiable	Tamper-evident authorization artifacts; evidence sufficient for independent replay	Governance demonstrated by record, not assertion

Level 5: Independently Verifiable Governance

At Level 5, the contractor is designed to produce tamper-evident authorization artifacts for consequential AI actions, with evidence sufficient for independent replay. The maturity endpoint is not better monitoring. It is an independently reconstructable verdict: whether the action was authorized before it executed.

OBSERVABILITY IS NOT AUTHORIZATION

Monitoring, logging, and after-the-fact review show that an AI system did something. They do not establish that the effect-bearing action was permitted before it released. The governing control is the authorization boundary: the point at which the action is allowed to proceed, recorded in a tamper-evident authorization artifact. If that artifact cannot be produced, the action fails closed.
The authorization boundary is the point where an output becomes an effect-bearing action, and where evidence, not explanations, must exist. Level 5 is where that evidence becomes independently verifiable.

MONITORING INPUTS ARE NOT THE AUTHORIZATION BOUNDARY

Monitoring, logging, alerting, and observability tooling may support contractor oversight, incident response, audit readiness, and performance management. Bundling those functions does not create the governing release control. A contractor should be able to show that each effect-bearing action was authorized before release, not merely that it was observed or logged.

TECHNICAL FOUNDATION

The Five Tests Standard

The guides in this series rest on the Five Tests Standard (5TS), an open specification for what it means for an AI action to be governed. The tests apply directly to the contractor evidence packages and oversight procedures described above.

TEST	REQUIREMENT
Stop	The system can be halted before side effects occur.
Ownership	Each consequential decision maps to a named accountable authority.
Replay	The decision can be reproduced at the boundary.
Escalation	Control transfers at defined policy boundaries.
Provenance	The inputs grounding a verdict have an established origin. Origin, not truth.

The verdict space is ALLOW, DENY, and ABSTAIN. An ABSTAIN blocks execution pending authorized human override; escalation is the consequence of ABSTAIN, not a fourth verdict. Provenance is a test, not a verdict.

OPEN SPECIFICATION

The Five Tests Standard is published as an open, vendor-neutral specification.
Five Tests Standard (5TS) · Open standard DOI: 10.5281/zenodo.21040295 · Explanatory foundation DOI: 10.5281/zenodo.21048661
The standard is independent of any single implementation.

APPENDICES

Templates and Reference

APPENDIX A

Evidence Package Checklist

Use this checklist to verify evidence package completeness before government submission. This is the Tier 1 evidence package, to be complete before deployment or submission.

☐ AI System description with intended use and limitations

☐ Governed Unit definition: use case, workflow, and decision authority

☐ Tier classification with rationale and trigger assessment

☐ Training data documentation: sources, demographics, selection, limitations

☐ Validation report with performance metrics overall and by subgroup

☐ Civil-rights, safety, and fairness testing methodology and results

☐ Human oversight SOP: roles, authority, training, time allocation

☐ Incident log: complete history, not "no incidents"

☐ Monitoring procedures with metrics, thresholds, and alert process

☐ Authorization artifact evidence for effect-bearing actions

☐ Subcontractor evidence packages, if applicable

☐ All documents current, updated within the required timeframe

APPENDIX B

Contract Clause Response Guide

When you see these clauses in solicitations or contracts, here is what to prepare.

CLAUSE TYPE	WHAT IT REQUIRES	WHAT YOU NEED READY
AI System Definition	Scope reaching primary-purpose and integrated AI; disclosure of use	Complete AI inventory; disclosure process
Governed Unit	Each use case governed independently	Per-use-case documentation, not one model
Tier Classification	Risk-based classification; government can reclassify	Honest self-assessment; Tier 1 readiness for borderline cases
Evidence Package	Training data, validation, bias testing, oversight, incidents	Pre-built packages; response within the contract window

CLAUSE TYPE	WHAT IT REQUIRES	WHAT YOU NEED READY
Verification Rights	Pre-award testing; access to docs, personnel, test environments	Organized repository; identified SMEs; test environment access
Suspension Authority	Government can stop AI pending investigation	Incident response plan; rollback capability; rapid remediation
Flow-Down	Requirements flow to subcontractors	Standard sub provisions; vendor evidence packages on file

APPENDIX C

Proposal Differentiator Language

Sample language to substantiate governance readiness in proposals. Customize bracketed items. Use only claims that are true, documented, and supportable with artifacts available for government review.

AI Governance Framework

"[Company] maintains an enterprise AI governance framework aligned with OMB M-25-21 and NIST AI RMF. Our framework includes tiered classification of AI systems, standardized evidence packages, continuous monitoring, and incident response procedures. For this requirement, we have pre-classified [system] as Tier [X] and can provide the complete evidence package within the contract's specified response window."

Evidence Package Readiness

"Our standard AI evidence package includes training data provenance documentation, validation results with demographic subgroup analysis, bias testing methodology and results, human oversight procedures, incident history, and monitoring procedures. A sample evidence package from a comparable engagement, with sensitive content redacted, is available to demonstrate our documentation standards."

Subcontractor Management

"Our subcontract template includes AI governance flow-down requirements and audit rights. We maintain evidence packages from all AI subcontractors and can provide government access to subcontractor documentation as required. Our proposed [subcontractor] has confirmed the ability to meet Tier [X] evidence requirements."

Audit Readiness

"[Company] maintains audit-ready documentation for AI systems on federal contracts. Where applicable and documented, we can provide examples of prior audit support, evidence-package delivery, and remediation history."

APPENDIX D

Proposal Pack Insert

Insert this section into proposals responding to AI governance requirements. Customize bracketed items. Use only claims that are true, documented, and supportable with artifacts available for government review.

AI Governance Approach

"[Company] delivers AI capabilities with governance infrastructure that meets federal requirements at award. Our approach is built on OMB M-25-21 alignment, NIST AI RMF principles, and evidence-based verification rather than assertion. We classify each AI use case as a Governed Unit with independent documentation, maintain pre-built evidence packages for rapid deployment, and operate continuous monitoring with defined escalation paths. For this requirement, we have assessed [system or capability] as Tier [1, 2, or 3] based on [rationale] and can deliver the complete evidence package within [30] days of award. Where applicable and documented, [Company] can provide examples of prior federal AI governance support, evidence-package delivery, audit response, and remediation history."

Requirement-to-Artifact Mapping

GOVERNMENT REQUIREMENT	OUR ARTIFACT	DELIVERY
AI System inventory and disclosure	Governed Unit Registry	Day 1
Tier classification with rationale	Tier Classification Memo	Day 1
Training data documentation	Training Data Provenance Report	Day 30
Validation with subgroup analysis	Validation and Civil Rights, Safety, and Fairness Testing Report	Day 30
Human oversight procedures	Human Oversight SOP	Day 30
Monitoring and drift detection	Monitoring and Drift Plan	Day 30
Incident response	Incident Response Playbook	Day 1
Subcontractor governance	Subcontractor Flow-Down Addendum	Day 1

AI Governance Deliverables

- **Tier 1 Evidence Package,** complete governance documentation per the applicable Section H requirements.
- **Governed Unit Registry,** an inventory of all AI systems with tier classification and rationale.
- **Monitoring and Drift Plan,** with metrics, thresholds, alert procedures, and review cadence.
- **Incident Response Playbook,** with detection, escalation, notification, and remediation procedures.
- **Subcontractor Flow-Down Addendum,** with standard AI governance provisions for all AI subcontracts.
- **Quarterly Governance Report,** with status, metrics, incidents, and remediation summary.
- **Annual Revalidation Report,** a full validation refresh for Tier 1 systems.

APPENDIX E

Evidence Package Table of Contents

Standardize your evidence packages with this structure. Government reviewers evaluate completeness, currency, specificity, and traceability. This is the standard table of contents for a Tier 1 evidence package.

Section 1: Executive Summary

System identification and Governed Unit definition; tier classification with rationale and trigger assessment; package version and currency statement.

Section 2: System Description

Architecture and technical approach; intended use and decision authority boundaries; known limitations and contraindications.

Section 3: Training Data Documentation

Data sources and provenance chain; demographic composition and representativeness analysis; selection criteria and preprocessing; known gaps and limitations.

Section 4: Validation Report

Validation methodology and test dataset description; performance metrics overall; performance metrics by demographic subgroup; civil-rights, safety, and fairness testing methodology and results; identified limitations and operating constraints.

Section 5: Human Oversight

Reviewer roles, authority, and accountability; training requirements and records; time allocation and workload analysis; escalation procedures.

Section 6: Monitoring and Operations

Metrics tracked and thresholds; drift detection approach; alert and escalation procedures; review cadence.

Section 7: Incident History

Incident log, complete; root cause analyses; corrective actions and effectiveness verification.

Section 8: Subcontractor Documentation, if applicable

Subcontractor AI inventory; evidence packages from each subcontractor; flow-down verification.

APPENDIX F

Credible Delivery Timeline

Use this timeline to make realistic commitments. Overpromising costs credibility; underpromising loses contracts. This is what governance-ready contractors can actually deliver.

Day 1 Deliverables, at Contract Award

DELIVERABLE	CONTENTS	PRE-CONDITION
Governed Unit Registry	Complete inventory of AI with tier classifications	Pre-built from proposal analysis
Tier Classification Memo	Rationale for each classification, trigger assessment	Completed during proposal
Incident Response Playbook	Detection, escalation, notification procedures	Standard template customized
Subcontractor Flow-Down Addendum	Standard AI governance provisions	Pre-negotiated with subs
Governance POC Designation	Named individuals with contact information	Staffing plan confirmed

Day 30 Deliverables

DELIVERABLE	CONTENTS	EFFORT REQUIRED
Training Data Documentation	Provenance, demographics, limitations	Compile from existing records plus gap analysis
Validation Report, initial	Performance metrics, subgroup analysis	Run validation suite on deployed configuration
Human Oversight SOP	Roles, authority, training, time allocation	Customize standard template to the contract
Monitoring and Drift Plan	Metrics, thresholds, alerting	Configure monitoring for the specific deployment

Day 90 Deliverables

DELIVERABLE	CONTENTS	DEPENDENCY
Complete Evidence Package	All sections populated, reviewed, approved	Day 30 deliverables plus operational experience
Civil Rights, Safety, and Fairness Deep Dive	Extended testing across all relevant subgroups	Sufficient operational data
First Quarterly Report	Metrics, incidents, governance status	Ninety days of monitoring data
Subcontractor Evidence Packages	Complete packages from all AI subs	Sub delivery plus your verification

Proposal guidance. Commit to Day 1 and Day 30 deliverables in your proposal. Reference Day 90 as the complete evidence package milestone. Do not promise Day 1 delivery of items that genuinely require

operational data.

APPENDIX G

What Not to Claim

These claims cost credibility with sophisticated government reviewers. Avoid them.

DO NOT CLAIM	WHY IT BACKFIRES
"100% accuracy" or "no false positives"	All AI has error rates. Claiming perfection signals either a misunderstanding of AI or hidden limitations.
"The vendor handles governance"	You face the government, not the vendor. This signals you cannot produce evidence when asked.
"It is not AI, it is analytics"	Contract definitions reach effect-bearing AI. This signals an attempt to evade governance requirements.
"No incidents to report"	Empty incident logs suggest problems are not being detected. Document issues with resolutions.
"Fully explainable AI," without qualification	Explainability is a spectrum. Specify what level you provide and for which decisions.
"Human in the loop," without specifics	Checkbox approvals do not count. Specify authority, training, time allocation, and accountability.
"Proprietary methodology," as a reason for no documentation	The government has audit rights. "Proprietary" is not a shield against evidence requirements.
"Tier 3" for anything citizen-facing or rights-affecting	The government can reclassify. Under-tiering will be caught and damages credibility.

Demonstrate mature governance by documenting limitations honestly, quantifying error rates, and showing how you manage the risks you have identified. Credibility comes from realism, not perfection.