MITRE

ETI EMERGING TECHNOLOGIES INSTITUTE

NDIA

Keeping Humans in Command

Preserving Human Agency in AI-Driven Federal Acquisition

NDIA Emerging Technologies Institute
The MITRE Corporation
May 2026

The research presented in this report was supported by the Acquisition Research Program, Graduate School of Defense Management at the Naval Postgraduate School.

To request defense acquisition research, please contact:

Acquisition Research Program
Department of Defense Management
Naval Postgraduate School
E: arp@nps.edu
www.acquisitionresearch.net

Copies of Symposium Proceedings and Presentations; and Acquisition Sponsored Faculty and Student Research Reports and Posters may be printed from the **NPS Defense Acquisition & Innovation Repository** at **https://dair.nps.edu/**.

ACQUISITION RESEARCH PROGRAM
DEPARTMENT OF ACQUISITION, FINANCE, AND MANPOWER
NAVAL POSTGRADUATE SCHOOL

Keeping Humans in Command:
Preserving Human Agency in AI-Driven Federal Acquisition

Stephen W. Roe—is a Managing Director at MITRE, leading Veterans and Military Health Modernization and supporting VA, VACO, and VHA since 2020. With over 20 years of experience driving cross-agency innovation, he has delivered measurable outcomes across healthcare, acquisition reform, and data-driven policy. His achievements include the Whole Health Recognition Framework (2+ million Veterans), PolicyNet (40% reduction, $55 million savings), and GenHub (2,000+ studies, 1,500 publications). He has shaped major VA priorities such as the PACT Act, EHR modernization, and community care reform, impacting $100+ billion in procurement while advancing AI-enabled patient safety and connected care. He holds degrees from Virginia Tech, Johns Hopkins, and UVA Darden. [sroe@mitre.org]

Dr. Nathan Turnipseed—serves as Director of the Acquisition Integration Service within the Department of Veterans Affairs Strategic Acquisition Center, where he provides strategic counsel to senior executives on enterprise acquisition life-cycle management, data governance, and complex acquisition strategy across a multi-billion-dollar portfolio. A veteran with more than 25 years of leadership in public and private sector acquisition environments, he is recognized for advancing large-scale, technology-enabled transformation initiatives and driving organizational performance. His distinguished U.S. Air Force career included leadership in personnel development for enlisted Airmen and commissioned officers. Dr. Turnipseed holds a Doctor of Management in Organizational Leadership, a Master of Management, and a Bachelor of Science in Psychology and Sociology, and maintains a Level III Federal Acquisition Certification for Program and Project Managers (FAC-P/PM). [Nathan.Turnipseed2@va.gov]

Ryan Novak—Acquisition Outcome Lead at MITRE, brings 30 years of experience as a USAF Contracting Officer and acquisition innovator. He holds multiple advanced degrees, including two MBAs, and is DAWIA III Contracting certified. A published author and speaker, Ryan created the Challenge-Based Acquisition Handbook and taught innovative acquisition approaches to 800+ professionals across government, industry, and MITRE. He has led strategic programs such as iCAMS, held an unlimited CO warrant, and completed executive programs at Oxford and Yale. As MITRE's Acquisition Innovation Center Lead for AI in Acquisition, he drives AI strategies, tools, and prototypes to transform federal procurement under the ACQAIRE Portfolio. [rnovak@mitre.org]

Adam Bouffard—is a Group Leader and Principal Decision Analyst at MITRE, with over 20 years of experience in acquisition and contracts supporting multiple DoD and civilian government agencies. Prior to joining MITRE in 2015, he worked as a civilian supporting the Naval Sea Systems Command and the Office of Naval Research. He holds a master's degree in systems engineering from George Washington University and possesses multiple contracts certifications. [abouffard@mitre.org]

Wilson Miles—is an Associate Research Fellow at ETI. His research and analysis portfolio focuses on emerging technology supply chains, acquisition policy, artificial intelligence, workforce issues, and other modernization technology policy issues. Wilson previously interned at several nonprofit organizations, including CRDF Global, the Hudson Institute, and the Foundation for Defense of Democracies. He earned his master's degree in international affairs: U.S. Foreign Policy and National Security from American University's School of International Service and a bachelor's degree in international relations from Linfield University. [wmiles@ndia.org]

Adam Kitay—is a Strategic Business Operations Principal at MITRE, with over 20 years of experience in government contracts and acquisition supporting various intelligence, DoD, and civilian agencies. Prior to joining MITRE, Adam worked as a civilian for the Department of the Navy, Naval Air Warfare Center (JSF PMO), the Environmental Protection Agency, and the Missile Defense Agency. He was also a Sr. Manager at Lockheed Martin Space, working in their Special Programs division. Adam holds a Bachelor of Science in Business Administration from Colorado State University. [akitay@mitre.org]

Kevin Forbes—is a Principal Software Engineer at the MITRE Corporation, with over 25 years of software development experience. His current work centers on applying Generative AI to tackle challenges in government acquisition and the software development life cycle. Kevin earned his

bachelor's degree in computer science from Georgetown University and a master's degree in systems engineering from Johns Hopkins University. [forbes@mitre.org]

Christopher Barlow—is an Innovation Strategist with The Belmont Group, with 10+ years of experience and demonstrated expertise in audit risk mitigation, acquisition data analytics, process development and optimization, and innovative acquisition approaches, including challenge-based acquisitions (ChBA) and artificial intelligence applications. His diverse experience spans multiple federal agencies. Chris holds a bachelor's degree in business administration from the University of Pikeville and is a Certified Business Intelligence Professional from Villanova University [cbarlow@blmtgrp.com]

Abstract

As artificial intelligence (AI) rapidly transforms government operations and private industry, the federal acquisition process faces a new paradox: the increasing use of AI to generate, evaluate, and award contract proposals. This convergence creates an "elephant in the room" scenario, where AI-enabled procurement systems in government may ultimately review proposals drafted by AI tools in industry, potentially diminishing the role of human judgment in critical decision-making processes. While automation offers efficiency, consistency, and scalability, the absence of human oversight risks eroding accountability, fairness, ethical discernment, and institutional knowledge in the acquisition process.

This paper examines the dual-use evolution of AI in federal procurement, specifically how government acquisition offices are adopting AI-driven tools for market research, requirements generation, and proposal evaluation, while industry leverages similar technologies to optimize proposal writing and cost modeling. Drawing on guidance from America's AI Action Plan and other federal AI policy frameworks, this research identifies the core risks of an AI-to-AI procurement ecosystem: diminished transparency, amplified algorithmic bias, and the potential loss of critical human judgment in evaluating qualitative and context-sensitive factors.

The expected outcome of this research is a framework for preserving meaningful human oversight within AI-augmented acquisition environments. This includes recommendations for: (1) defining human-in-the-loop thresholds across acquisition phases; (2) establishing verification and audit mechanisms for AI-generated content; and (3) integrating AI literacy training for acquisition professionals.

By confronting the AI-to-AI dilemma directly, this study seeks to confirm that automation enhances, rather than replaces, the human expertise, ethical reasoning, and strategic judgment that underpin the integrity of federal acquisitions. The goal is a balanced approach where AI supports data-driven decision-making, while humans remain the final arbiters of fairness, accountability, and mission alignment, preserving the essential role of human agency at the core of accelerating warfighting capabilities. This research directly aligns with the theme of this year's Naval Postgraduate School Acquisition Research Symposium, "Accelerating Warfighting Capabilities," by advancing an ethical and operational framework that confirms AI adoption strengthens, rather than undermines, the speed, trust, and integrity of defense acquisition.

Introduction

Artificial intelligence (AI) is rapidly reshaping the federal acquisition environment. Government acquisition organizations are beginning to deploy AI-enabled tools to assist with market research, requirements development, cost analysis, and proposal evaluation. At the same time, contractors across the defense industrial base are adopting generative AI and advanced analytics to automate proposal writing, compliance checks, and pricing optimization. These parallel developments are converging within the same procurement ecosystem, fundamentally altering how acquisition information is produced, analyzed, and acted upon.

This convergence introduces a structural governance challenge for federal procurement. When government AI systems evaluate proposals that have been drafted or optimized by AI tools, acquisition decisions may increasingly be shaped by machine-mediated interactions

rather than expert human judgment. Although automation promises significant gains in speed, consistency, and analytical scale, capabilities strongly aligned with the Department of War's imperative to accelerate warfighting capability delivery, it also introduces new risks related to transparency, bias amplification, and accountability.

This paper conceptualizes this emerging dynamic as the "AI-to-AI procurement paradox." In such an environment, automated systems on both sides of the acquisition process may influence information inputs, proposal structures, and evaluation outcomes before human decision-makers can intervene. Without deliberate governance mechanisms, this dynamic could erode human reasoning, ethical oversight, accountability, and subject matter expertise that traditionally anchor federal procurement decisions.

"AI IS RAPIDLY CHANGING OUR WORLD AND HAS SIGNIFICANT POTENTIAL TO TRANSFORM SOCIETY AND PEOPLE'S LIVES. HOWEVER, AI TECHNOLOGIES ALSO POSE RISKS THAT CAN NEGATIVELY IMPACT INDIVIDUALS, GROUPS, ORGANIZATIONS, COMMUNITIES, SOCIETY, AND ENVIRONMENTS." (GAO, 2025)

The Rise of AI in Federal Acquisition

Federal acquisition has historically adapted to technological advancements to improve efficiency and effectiveness, evolving from paper-based contracting to fully digitized procurement systems. AI represents the next significant evolution in this trajectory. Acquisition organizations are now experimenting with or deploying AI-enabled tools to analyze market intelligence, assist in drafting requirements, support pricing and cost realism assessments, and augment proposal evaluation processes.

These capabilities align closely with contemporary defense acquisition priorities, particularly the imperative to accelerate capability delivery in response to rapidly evolving threats. AI offers the ability to process large volumes of data at speeds unattainable by human analysts, identify trends across complex datasets, and promote consistency in evaluative decision-making. From an operational perspective, such efficiencies are attractive in an environment characterized by constrained resources, reductions in the size and experience of the acquisition and engineering workforces of the Department of War, and increasing demand for rapid acquisition outcomes.

However, the integration of AI into acquisition processes is also reshaping decision-making authority. As AI systems provide increasingly sophisticated recommendations, there is a growing risk that human decision-makers may defer to algorithmic outputs, especially under time pressure. This shift underscores the need to examine not only how AI is used, but also how it alters the balance between automation and human judgment in acquisition governance. iQuasar (2026) summarizes the underlying dilemma: "AI procurement in 2026 will transform federal contracting through automated evaluations, risk monitoring, and data-driven oversight." This anticipated transformation highlights the importance of understanding both the benefits and the challenges AI brings to acquisition governance, setting the stage for further discussion of its broader implications.

The AI-to-AI Dilemma

The rapid adoption of AI within government acquisition is occurring alongside significant advancements in AI-enabled proposal development within industry. The defense industrial base is leveraging AI tools to automate compliance checks, generate technical narratives, optimize pricing strategies, and tailor proposals to solicitation requirements. These tools increase competitiveness and efficiency, but they also fundamentally change the nature of proposal

inputs entering the federal acquisition system.

The convergence of these trends creates an acquisition environment where automated government systems may increasingly assess, score, or filter proposals generated by automated industry systems. In this scenario, human involvement may be limited to overseeing machine-generated outputs rather than making substantive evaluative judgments. This dynamic raises several critical risks: transparency may be diminished if decision logic is embedded within opaque algorithms; algorithmic bias may be amplified if both proposal generation and evaluation rely on historically biased data; and accountability may be weakened, complicating responsibility for acquisition outcomes.

Most importantly, the AI-to-AI paradox threatens to erode the human judgment necessary for evaluating qualitative and context-sensitive factors such as innovation, operational risk, and mission alignment. These factors are central to defense acquisition success yet are difficult to fully capture through automated analysis. Without intentional safeguards and frameworks to navigate this dilemma, the pursuit of speed and efficiency may inadvertently undermine the integrity and trust that sustain effective acquisition.

Purpose and Thesis

The purpose of this paper is to examine the AI-to-AI procurement paradox and assess its implications for federal acquisition integrity and warfighting capability delivery. By analyzing the parallel adoption of AI by government and industry, this research identifies the ethical, operational, and governance risks that emerge when automation increasingly mediates acquisition decisions. As noted, “Agencies still need a human in the loop ... but that person is reviewing something created instantly, instead of starting it from scratch manually or spending time filling in details in a blank template” (FNN, 2025). This shift underscores the importance of maintaining meaningful human oversight even as automation accelerates the acquisition process.

The central thesis of this paper is that while AI has the potential to significantly enhance acquisition speed and analytical rigor, it must not replace human judgment in critical decision-making processes. FAR 7.503(c)(12) explicitly classifies these responsibilities as inherently governmental, requiring direct human-in-the-loop involvement, and recent protest decisions reinforce that the source selection official must perform an independent assessment beyond the work of agency evaluators before making a final decision. Accordingly, AI should be deliberately integrated into a human-centered acquisition framework that preserves accountability, confirms ethical reasoning, and maintains sound strategic judgment. To address this challenge, this paper proposes a governance model called the Human Agency Assurance Framework (HAAF), designed to preserve meaningful human control within AI-enabled acquisition environments. The framework integrates oversight thresholds, verification and audit mechanisms, explainability requirements, governance structures, and workforce competencies to confirm that AI augments rather than replaces expert acquisition judgment.

Background: Guidance and Precedent

The federal government has acknowledged the transformative potential of AI while recognizing the governance challenges it presents. Through various directives, frameworks, and guidance, it has sought to establish a foundation for responsible AI development and deployment. However, despite these efforts, risks associated with the use of AI for acquisition-specific functions, including procurement, integration, and life-cycle management, lack robust federal policy to regulate and optimize their use.

At the national level, America’s AI Action Plan underscores the importance of ensuring AI systems are ethical, transparent, and accountable (The White House, 2025). It highlights the

need for governance frameworks that address issues such as bias, privacy, and security. While the plan promotes innovation and public-private collaboration, it does not explicitly address the risks tied to the acquisition and operationalization of AI systems, particularly in complex environments like defense and national security.

The DOW has taken a more targeted approach to AI governance through its Ethical Principles for AI, which emphasize responsible, equitable, traceable, reliable, and governable AI (DOW, 2026). These principles are intended to guide the development and deployment of AI in defense applications. DOW's AI Strategy further outlines the need for robust testing, evaluation, and validation of AI systems to confirm they meet mission-critical requirements. However, gaps remain in addressing risks to the acquisition of AI systems, such as supply chain vulnerabilities, vendor accountability, and the long-term sustainment of AI technologies.

From an administrative perspective, the Office of Management and Budget (OMB) has issued guidance on the use of AI in federal agencies, including the OMB Memorandum M-21-06, which provides principles for AI use in government (GSA, n.d.). This guidance emphasizes transparency, accountability, and risk management but, like the documents referenced above, it does not delve into the unique challenges of acquiring AI systems, such as ensuring compliance with ethical standards during procurement or managing the risks of proprietary technologies.

Finally, Executive Order 14110 highlights the federal government's commitment to advancing responsible AI (GSA, 2025). It calls for interagency collaboration and the establishment of standards to govern AI use. However, like other federal directives, it does not specifically address the acquisition life cycle, leaving a critical gap in governance.

In summary, while federal directives, DOW frameworks, and administrative guidance collectively demonstrate a commitment to responsible AI governance, they do not fully address acquisition-specific risks. These risks include ensuring regulatory and ethical compliance during procurement, managing vendor relationships, mitigating supply chain vulnerabilities, and sustaining AI systems over time. This gap highlights the need for more comprehensive policies that integrate acquisition-specific considerations into the broader AI governance framework. Developing agency-specific Human Agency Assured Frameworks (HAAF) tailored to your organization is therefore a critical step, emphasizing the importance of human agency.

Precedent: Why Human Agency Matters in Acquisition

Human agency is essential in the AI acquisition life cycle because it confirms that human judgment, oversight, and accountability remain central throughout procurement, deployment, and life-cycle management. In the context of acquisition, human agency is important for several key reasons, as shown in Table 1.

Table 1. The Role Human Agency in Acquisition

Key Area	Description
Ethical Decision Making	Human agency ensures that decisions about the selection, design, and deployment of AI systems align with ethical principles, such as fairness, accountability, and transparency.
Accountability and Responsibility	Human agency ensures that individuals or organizations remain accountable for the outcomes of AI systems.
Risk Mitigation	The use of AI systems involves inherent risks, such as supply chain vulnerabilities, cybersecurity threats, and potential misuse of technology. Human agency is necessary to identify, assess, and mitigate these risks throughout the acquisition process.
Contextual Understanding	AI systems often lack the ability to fully understand the context in which they operate. Human institutional expertise bridges the gap between AI outputs and hallucinations.
Trust and Transparency	Trust is a cornerstone of successful AI adoption, and human agency plays a vital role in building and maintaining that trust. When humans are actively involved in the acquisition process, they can ensure that AI systems are transparent, explainable, and aligned with stakeholder expectations.
Adaptability and Oversight	AI systems are not static; they evolve over time through updates, retraining, or changes in operational environments. Human agency ensures that these changes are monitored and managed responsibly.
Avoiding Over Reliance on Automation	Over reliance on AI systems without human oversight can lead to automation bias, where humans defer to AI decisions even when they are flawed.

Scholars examining AI governance emphasize that maintaining human responsibility must begin at the earliest stages of system development and acquisition. Bode and Huelss (2025) state, "Retaining human responsibility in the development and use of AI-based systems requires establishing accountability mechanisms at the earliest stages of the lifecycle, ensuring that human agency is exercised across design, deployment, and oversight to prevent the diffusion of moral responsibility." Human agency in acquisition is essential to confirm that AI systems are procured, deployed, and managed in alignment with ethical principles, risk mitigation, and accountability. It confirms that humans remain in control of critical decisions, fostering trust, transparency, and adaptability throughout the life cycle of AI systems. Without human agency, the acquisition process risks becoming overly reliant on automation, which can lead to ethical, operational, and security challenges. Next, this paper will take a deeper dive into the significant risks that arise in an AI-to-AI procurement ecosystem.

Analysis: Risks of an AI-to-AI Procurement Ecosystem

"GIVEN THE RAPID GROWTH IN CAPABILITIES AND WIDESPREAD ADOPTION OF AI, THE FEDERAL GOVERNMENT HAS TAKEN ACTION … INCLUDING TO GUIDE AGENCIES IN ACCELERATING THE USE AND ACQUISITION OF AI." (GAO, 2025)

As federal acquisition organizations integrate AI into various stages of the acquisition life cycle, a set of interrelated risks emerges that extend beyond isolated technical concerns. These risks stem from how AI tools are embedded in the life cycle, governed across organizations, and relied upon by the workforce. In an AI-to-AI procurement ecosystem, these risks are compounded, resulting in systemic vulnerabilities that threaten transparency, competitive fairness, innovation, and alignment with the warfighting mission.

This section analyzes four primary risk categories: opaque reasoning and loss of transparency; algorithmic bias amplification; proposal homogenization and over optimization; and erosion of human judgment and contextual understanding. While analytically distinct, these

risks reinforce one another as AI adoption accelerates throughout all phases of the program and the entire acquisition life cycle.

Opaque Reasoning and Loss of Transparency

A foundational principle of federal acquisition is the government's obligation to justify procurement decisions in a transparent, traceable, and defensible manner. Source selection outcomes must be defensible by contracting officers, oversight authorities, and, when necessary, judicial bodies. The introduction of AI-enabled tools into requirements development, evaluation, and decision support challenges this obligation when model outputs cannot be clearly explained or reproduced.

Many AI systems used for proposal scoring, risk management, or text generation operate as black boxes, with logic that is proprietary or opaque even to their developers (NIGP, 2026). These systems may produce precise rankings or narrative assessments without revealing which features were most influential or how trade-offs were resolved. When such outputs materially shape evaluations, acquisition officials may struggle to articulate why one offeror was rated superior, relying solely on the tool's conclusions. As Tillipman (2026) notes, "When agencies rely on AI systems whose conclusions cannot be meaningfully examined because the systems are opaque, technically complex, or contractually shielded, the government may be able to show what the tool produced, but not why those outputs reflect the reasoned judgment the APA requires." This opacity undermines legal defensibility under the Federal Acquisition Regulation (FAR), complicates debriefings, increases protest risk, and erodes confidence in the fairness of outcomes.

Transparency risks extend beyond evaluation to AI-assisted drafting of solicitations, requirements, and evaluation criteria. When AI tools generate or heavily influence acquisition documents, the government remains responsible for maintaining a defensible record of what inputs were used, which models were employed, how outputs were edited, and by whom. Weak logging, version control, or origin tracking complicates records management and post-award reviews. It also blurs the boundary between human judgment and machine recommendation, which is particularly problematic when ambiguities or errors in AI-generated language later become grounds for protest.

These challenges are magnified in high-consequence defense acquisitions. When AI-assisted analyses inform judgments about survivability, operational suitability, or system-of-systems integration, senior leaders and warfighters must understand the reasoning behind those assessments. The inability to clearly explain why a particular solution was deemed superior undermines confidence not only in the acquisition outcome, but also in the AI tools themselves.

Algorithmic Bias Amplification

Algorithmic bias is a critical ethical and acquisition risk in an AI-to-AI procurement ecosystem, especially because bias can be amplified rather than mitigated when AI systems interact throughout the life cycle. In acquisition, AI models learn from historical data that reflects past award decisions, evaluation practices, and industrial base structures. This data is generated and stored across distributed government and industry systems, including SAM.gov, CPARS, agency contract writing systems, and program office repositories, vendor proposals, and market research artifacts. When these historical patterns are embedded in both government evaluation tools and industry proposal generation systems, bias becomes self-reinforcing.

On the government side, AI-enabled market research or proposal triage tools may implicitly favor vendors, approaches, or language resembling past winners, often privileging incumbents and established primes. On the industry side, proposal generation tools trained on

prior successful submissions optimize content to match those same patterns. When evaluative AI systems rely too heavily on similarity to past examples, rather than considering past performance as just one factor, innovative or unconventional solutions can be unfairly disadvantaged.

This feedback loop has direct implications for competitive fairness. Small businesses, the nontraditional defense industrial base, and other new entrants may find it increasingly difficult to compete if AI-mediated processes implicitly favor scale, legacy performance profiles, or standardized proposal structures. Conversely, traditional defense businesses, experienced small businesses, federally funded research and development centers (FFRDCs), UARCs, and universities could also be disadvantaged if automated processes are biased against their proposals and toward those of new entrants. Even when explicit demographic variables are excluded, proxies such as company size, contract history, or terminology can reintroduce disparate impacts. Over time, AI-to-AI interactions risk narrowing the competitive landscape, reinforcing incumbency advantages, and undermining statutory objectives related to fair and open competition, while giving the appearance of objective, data-driven decision-making. U.S. DoS (2024) notes that harmful bias and homogenization result in "amplification and exacerbation of historical, societal, and systemic biases; performance disparities between sub-groups or languages ... undesired homogeneity in data inputs and outputs resulting in degraded quality of outputs." All of these factors complicate the overarching paradox that our proposed framework seeks to address.

Proposal Homogenization and Over Optimization

The widespread adoption of AI in proposal development introduces strong incentives for over-optimization. Generative AI tools are designed to maximize compliance with solicitation requirements, align language precisely with evaluation criteria, and minimize perceived risk. While these capabilities can improve efficiency and baseline proposal quality, they also encourage convergence around a narrow set of "model-approved" narratives and structures.

As multiple offerors rely on similar tools trained on overlapping data from past solicitations and winning proposals, technical and management volumes increasingly resemble one another. MeriTalk (2026) highlights "a growing risk that AI could harm the integrity of the acquisition process, citing an uptick in protests and filings that rely on AI-generated content, including fabricated cases and citations—a byproduct of ongoing concerns about hallucinations." Novel architectures, unconventional teaming strategies, or disruptive technologies may be normalized into familiar patterns or excluded entirely if they deviate from historical examples. When government evaluators also rely on AI-assisted tools optimized for pattern recognition, this homogenization is reinforced rather than challenged.

This dynamic poses a direct risk to innovation and long-term warfighting advantage. Kolt (2025) observes, "While the homogenization introduced by foundation models promotes efficiency ... it also gives rise to new risks familiar to complexity researchers." Defense acquisition depends on identifying solutions that are not only compliant, but also operationally effective, adaptable, resilient, and responsive to evolving threats. An acquisition environment that implicitly rewards conformity and linguistic optimization over substantive differentiation risks selecting low-variance solutions that appear safe on paper but lack strategic depth. Over time, this suppresses experimentation and reduces the government's exposure to transformative capabilities.

Loss of Human Judgment and Contextual Understanding

The most significant risk in an AI-to-AI procurement ecosystem is the erosion of human judgment. Federal acquisition is inherently context-dependent, requiring experienced professionals to balance technical performance, cost, risk, schedule, and mission relevance.

However, AI tools introduce strong incentives for cognitive offloading, particularly under time pressure and resource constraints.

Automation bias can lead evaluators to over trust AI-generated scores, summaries, or narratives, especially when these are presented with quantitative precision or authoritative language. Over time, acquisition personnel may shift from exercising independent judgment to validating algorithmic outputs. This is particularly problematic for qualitative assessments such as innovation potential, management credibility, teaming viability, and adaptability to evolving operational needs, areas where context and experience are essential.

There is also a long-term institutional risk. If AI routinely drafts requirements, evaluation narratives, or acquisition strategies, newer professionals may never fully develop the tacit expertise required to detect unrealistic assumptions, hidden execution risks, or subtle misalignments with warfighter needs. AI systems trained primarily on historical data (e.g., prior RFPs and SOWs, source selection evaluation reports, CPARS past performance records, archived vendor proposals, etc.) cannot fully account for emerging concepts of operations, evolving adversary tactics, or cross-domain dependencies. Without deliberate human-in-the-loop engagement, acquisition decisions may optimize for what is easiest to measure rather than what is most critical to mission success. Additionally, transferring the institutional knowledge of subject matter experts with years of experience is extremely difficult, potentially leaving a significant gap in the AI's inputs.

This tension between analytical efficiency and human judgment is widely recognized in the broader AI governance literature. "The key challenge is to harness AI's computational capabilities in crisis decision-making while preserving human moral agency. The goal is not to replace human decision-makers, but rather to create a collaborative interface where AI can support and enhance ethical reasoning without undermining fundamental moral responsibility" (Johnson, 2026). Over time, acquisition personnel may shift from exercising independent judgment to merely validating algorithmic outputs.

Compounding Effects Across the Acquisition Life Cycle

Individually, each of these risks is manageable. Collectively, however, they form a compounding threat to acquisition integrity. Opaque reasoning obscures accountability; bias amplification distorts competition; proposal homogenization suppresses innovation; and diminished human judgment weakens mission alignment. As AI is applied sequentially, from requirements framing, to proposal generation, to evaluation, these dynamics reinforce one another, creating systemic vulnerabilities that are difficult to detect and correct after the fact.

Addressing the AI-to-AI procurement paradox therefore requires more than technical safeguards applied at isolated points in the process. As the Open Contracting Partnership (2025) states, poor AI use in procurement "can cost governments millions, not just in failed projects, but in lost public trust." Addressing these risks demands a holistic, life-cycle-oriented approach that preserves transparency, fairness, and human agency as AI becomes embedded in federal acquisition. The following section builds on this analysis by proposing a framework for maintaining meaningful human oversight within AI-augmented procurement environments.

Proposed Framework and Recommendations: Preserving Human Agency in AI-Enabled Acquisition

Federal acquisition organizations are rapidly integrating AI into all aspects of the acquisition life cycle. While these applications can reduce lead times and enhance analytical capabilities, they also introduce new failure modes. Recent federal direction emphasizes risk management, transparency, and accountability practices for AI, especially for uses that impact public rights and safety (OMB, 2024).

Proposed Human Agency Assurance Framework (HAAF)

The proposed HAAF pairs acquisition decision authority with measurable AI assurance gates. It is designed to be implemented within existing acquisition governance structures (e.g., source selection plans, quality assurance surveillance plans, and risk registers) and aligns with federal AI governance expectations for inventorying AI use cases, managing risks, and documenting controls (OMB, 2024). Figure 1 illustrates the core principles of the proposed HAAF from a source selection perspective.

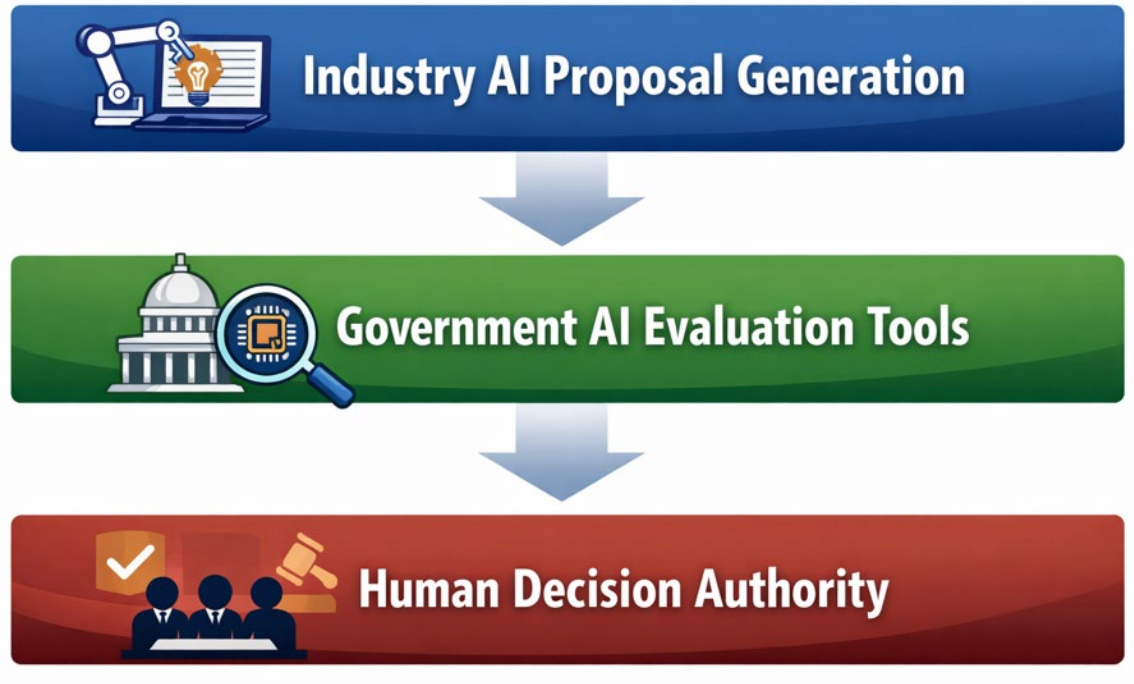

Figure. 1 HAAF for Source Selection

HAAF consists of five integrated components, as shown in Table 2.

Table 2. Components of the HAAF

Component	Mechanism	Purpose
Oversight Thresholds	Human-in-the-Loop (HITL) and Human-on-the-Loop (HOTL) oversight tiers that define when humans must approve decisions, supervise automated processes, or directly execute actions.	Confirms that the level of human control scales with operational risk, mission impact, and ethical considerations, preserving meaningful human judgment in consequential decisions.
Assurance Gates	Verification, Validation, and Accreditation (VV&A) processes are combined with audit and traceability requirements throughout system development and deployment.	Provides evidence that AI systems function as intended, maintain transparency in decision logic, and meet reliability, safety, and compliance standards prior to operational use.
Governance	Provides evidence that AI systems function as intended, maintain transparency in decision logic, and meet reliability, safety, and compliance standards prior to operational use.	Establishes institutional controls to prevent misuse, confirm responsible deployment, and maintain alignment with organizational objectives and ethical standards.
Policy and Guidance	Formal AI policies, acquisition guidance, and implementation frameworks that translate strategic objectives into operational rules and procedures.	Provides consistent standards for integrating AI into acquisition and operational environments, ensuring compliance with legal, ethical, and regulatory requirements.
Workforce Readiness	Role-based AI literacy training, competencies, and operating procedures for acquisition professionals, engineers, and oversight personnel.	Sustains accountability and informed decision-making by equipping the workforce with the knowledge required to evaluate, supervise, and govern AI-enabled systems.

Before examining the five components, it is essential to emphasize that the success of HAAF depends on data quality. This requires a clear understanding of the intended tool, the use of relevant and high-quality data, and sound model development. Aligning data with desired outcomes from the outset confirms the model is effective before training even begins. With that foundation, the five HAAF components are outlined below.

Human-in-the-Loop and Human-on-the-Loop Thresholds

Preserving human agency begins by defining which acquisition decisions can be AI-assisted versus AI-automated and, critically, where humans must remain the accountable decision-makers. In acquisition, “human agency” means that contracting officers and designated acquisition officials retain meaningful control over outcomes, can understand and challenge AI supported recommendations, and can intervene before harm occurs. Practical implementation requires explicit thresholds tied to decision criticality, uncertainty, and potential harm. HAAF uses a four-tier oversight model, as shown in Table 3.

Table 3. Tiered Tool Levels

Tier	Oversight
Tier 0	Automation Prohibited
Tier 1	Human-in-the-Loop
Tier 2	Human-on-the-Loop
Tier 3	Administrative AI

Tier 0 (Automation prohibited): AI may support drafting or retrieval but cannot recommend or execute decisions (e.g., determinations affecting eligibility, responsibility, or human rights and safety). This tier is appropriate for high-stakes, high-impact use cases and mirrors the federal emphasis on heightened practices for uses that affect human rights and safety (OMB, 2024).

Tier 1 (HITL approval required): AI may generate recommendations (e.g., draft evaluation narratives, risk assessments, price realism indicators), but a human must review the rationale and explicitly approve the output before it is used. HITL is required when the decision has material consequence, when data quality is uncertain, or when the model is newly deployed or recently updated.

Tier 2 (HOTL supervision required): AI may execute bounded tasks with continuous monitoring and rapid intervention capability (e.g., automated compliance checks against solicitation instructions; anomaly detection that flags transactions). HOTL is appropriate when the task is repetitive and low-consequence per instance, but aggregate errors could create material risk.

Tier 3 (Human-out-of-the-loop (HOOTL) allowed only for non-material administrative support): AI may operate with minimal oversight where errors are easy to detect and correct and have negligible impact (e.g., formatting, routing, de-duplication of non-decisional records).

Operationally, each AI-enabled acquisition should be assigned a tier by reviewing the following questions:

- Does the output affect award decisions, contractor selection, performance ratings, payment integrity, or rights/safety?
- Does the use of the AI systems significantly increase the likelihood of a protest of the award?
- If incorrect, can the output be corrected before harm occurs and without major cost/schedule/mission impact?
- Is the model operating outside the bounds of validated data (e.g., new commodity, new mission area, novel solicitation structure)?
- Can a reviewer replicate or meaningfully discern the basis for the recommendation?
- Is the acquisition exposed to intentional manipulation (e.g., prompt injection, data poisoning, strategic responses)?

There are two clear “tripwires” that require HITL oversight: (1) any use in source selection judgments (e.g., technical acceptability, best-value trade-offs, limiting eligible sources, past performance narratives, sole source justifications), and (2) any use that could influence enforcement or eligibility decisions. These controls complement federal accountability practices and reflect the acquisition community’s consensus that AI should augment, not replace, expert judgment in core decisions (MITRE, 2024). Acquisition practitioners increasingly describe AI as a decision support capability rather than a replacement for contracting authority. As Hubbard (2025) notes, “AI serves as a decision support tool, not a decision-maker. Acquisition professionals retain full authority over all procurement decisions, with AI providing analysis and recommendations that humans validate and approve.”

This can be achieved by requiring an AI Oversight Annex in acquisition and source selection plans that lists AI-enabled aspects, assigns the oversight tier, identifies accountable officials, and defines escalation triggers (e.g., confidence below threshold, drift indicators, team findings). This aligns with government-wide expectations for AI use case inventories and risk management documentation (OMB, 2024).

Verification, Validation, and Audit Requirements

To preserve agency, humans must be able to trust AI outputs but also verify them. VV&A should be scaled according to the oversight tier and associated risk. NIST’s Generative AI Profile for the AI Risk Management Framework outlines governance and measurement activities that help organizations identify and manage generative AI risks through repeatable evaluations (NIST, 2024). Similarly, OMB directs agencies to implement minimum risk management practices and documentation for AI uses, with heightened expectations for applications that impact human rights or safety (OMB, 2024).

The HAAF establishes four assurance gates that must be satisfied before an AI capability is used:

Gate A – Intended use definition and data pedigree: Document purpose, decision boundaries, prohibited uses, and data provenance. Include training data sources (or vendor attestations), sufficiency of available data, data rights, privacy constraints, and known limitations.

Gate B – Pre-deployment validation: Validate accuracy and robustness against representative acquisition artifacts (e.g., solicitations, proposals, market research summaries) and stress test for manipulation (e.g., prompt injection, instruction-following attacks). For scoring or ranking use cases, validate calibration and evaluate differential error rates across vendor types where feasible.

Gate C – Operational monitoring and drift management: Monitor performance over time (e.g., concept drift, data drift), track user overrides, and establish rollback criteria. Where AI influences decisions, log the “reason codes,” context, and evidence presented to the reviewer.

Gate D – Independent audit and accountability: Periodically conduct independent reviews of model performance, security controls, and documented decisions. GAO’s AI Accountability Framework emphasizes practices around governance, data, performance, and monitoring, and can serve as an audit checklist for federal AI deployments (GAO, 2024). GAO has also highlighted the rapid growth of generative AI use at agencies and the need for policy-compliant management amid fast technical change (GAO, 2025).

Auditability is particularly important for acquisition artifacts that may be sensitive to protests. The HAAF recommends that AI tools used in evaluation support produce defensible records, including versioned outputs, citations to underlying inputs, reviewer decisions (accept/modify/reject), and explanations of how the AI was used. This does not require full

model transparency for proprietary systems; rather, it requires decision traceability sufficient for oversight and after-action review. Legal scholars have similarly warned that agencies must preserve the ability to interrogate and override AI tools used in procurement decisions. As Tillipman (2026) notes, “Agencies that deploy AI tools should therefore secure contractual rights that preserve the conditions for independent judgment: sufficient information to understand how the tool applies the solicitation’s evaluation criteria and the practice ability to question, and where necessary, override or disregard its outputs.”

Require Explainability and Decision Traceability

AI-assisted evaluations must be transparent and interpretable. Explainability artifacts, including rationale summaries, confidence indicators, or input-output mappings, should accompany AI outputs and be retained in the administrative record. Transparency and interpretability are core elements of responsible AI governance (NIST, 2023). In acquisition contexts, explainability allows evaluators to demonstrate how AI informed decisions while ensuring that humans retain authority.

Explainability also strengthens protest defensibility by documenting the reasoning behind evaluation judgments. Without traceability, agencies risk relying on impervious outputs that cannot be justified under scrutiny.

AI Governance and Oversight

AI governance in acquisition should be treated as a system of controls encompassing people, process, and technology. OMB directs agencies to strengthen AI governance by establishing internal capabilities and implementing minimum risk management practices for certain AI uses (OMB, 2024). Complementary assurance guidance emphasizes continuous monitoring, documentation, and transparency regarding context and limitations to sustain accountability over time (NIST, 2024). In acquisition, governance must also preserve the integrity of the procurement record, ensuring that AI assistance does not introduce undocumented criteria, untraceable reasoning, or automation bias. Moreover, as demonstrated in the development of complex technical regulatory regimes, the best outcomes result from continuous and substantive dialogue among tool developers, acquisition personnel, and regulatory bodies.

AI governance includes ethical safeguards, which should be expressed as enforceable requirements rather than aspirational principles. The EU Artificial Intelligence Act (AI Act) adopts a risk-based approach with obligations for high-risk systems, offering a concrete template for procurement-facing controls (European Parliament and Council, 2024). For federal agencies, OMB’s direction similarly stresses risk management practices and governance structures for AI uses that may affect rights and safety (OMB, 2024). These considerations should be incorporated when formulating the governance and oversight associated with AI usage. The recommendations below are designed to be implementable through familiar acquisition governance artifacts and to scale across multiple AI use cases.

Establishing Cross-Functional AI Oversight Boards

Create a cross-functional AI Oversight Board that integrates acquisition, program management, legal, privacy, cybersecurity, data governance, and independent testing and audit functions. The board’s role is to set and maintain guardrails that confirm AI use is repeatable and defensible. At a minimum, the board should: maintain an inventory of AI-enabled acquisition use cases; assign oversight tiers and approval authorities by risk; approve validation plans and monitoring metrics; consult with industry on use of AI in acquisition processes; maintain awareness of state-of-the-practice AI capabilities applicable to acquisition activities; and review incidents, drift indicators, and material user overrides. This structure supports OMB’s emphasis on agency AI governance and aligns with accountability practices recommended by GAO for

federal AI use. To further enhance its impact, the board should publish a brief quarterly "AI acquisition governance bulletin" describing approved use cases, controls, and lessons learned to create institutional learning loops (GAO, 2025).

Embedding Governance in Contract Terms

When agencies procure AI tools or AI-enabled services, governance must be translated into enforceable contract language. Contracts should require contractors to support verification and monitoring activities, provide audit evidence, and notify the government of significant model changes. The government should use a concrete, risk-based set of obligations, including documentation, logging, human oversight, and robustness, for higher-risk systems to inform procurement requirements.

Minimum governance contract terms for AI acquisitions should include model and prompt configuration control and version disclosure, logging requirements and data retention for audit purposes, performance monitoring and drift notification, security testing for adversarial threats, and an obligation to support independent government assessments. These terms operationalize accountability and improve traceability of AI-assisted work products. Governance may also include a "model change clause" requiring advance notice and government approval for changes that could materially affect outputs used in acquisition decisions (e.g., evaluation support, fraud indicators, compliance screening).

Defining Qualitative Judgment Anchors

Acquisition activities often rely on qualitative judgments, such as best-value trade-offs, relevance of past performance, experience with previous acquisition decisions and their programmatic outcomes, or risk assessments. To preserve human agency, agencies should define qualitative judgment anchors—short, plain-language criteria that clarify what "good" looks like for common qualitative determinations and specify the evidence required to support them. Anchors reduce automation bias by requiring reviewers to map AI outputs to meaningful criteria, rather than accepting fluent text as a proxy for correctness.

Anchors can be incorporated into evaluation worksheets and Standard Operating Procedures (SOPs). For example, a past-performance narrative anchor might require the reviewer to cite the reference, provide a relevance rationale aligned with scope and complexity, and include corroborating record specifics. NIST's Generative AI profile emphasizes the importance of context, limitations, and measurement in managing generative risks (NIST, 2024). Anchors deliver this contextualization at the decision point. Additionally, the government could require an "AI-assisted work product checklist" for any AI-generated narrative used in the procurement record, confirming citations, factor alignment, and exclusion of non-evaluated criteria (GAO, 2025).

Requiring Industry to Disclose AI Use and Metadata

Human agency and fairness are compromised when the government cannot distinguish between vendor-authored content and AI-generated content, or when AI is used to optimize submissions in ways that conceal weaknesses or introduce unverifiable claims. Agencies should require offerors and contractors to disclose material AI use and associated metadata for acquisition-relevant artifacts (e.g., proposals, technical volumes, deliverable narratives, and automated test/analysis outputs).

A practical disclosure approach is to scope the requirement to material AI use, defined as either substantive content generation or automated analysis, and require disclosure of the tools used and their purpose; the tool and model version; the date, time, and high-level configuration or policy mode, if relevant; and a contractor attestation that all factual claims were human-verified. For high-dollar or higher-risk procurements, requiring an "AI content

provenance annex" that identifies AI-generated sections and documents human verification steps performed by the offeror would be ideal.

AI Policy and Guidance

Governance mechanisms are most effective when reinforced by clear policy and repeatable guidance. GAO (2025) has observed that agencies face challenges complying with evolving federal AI policies while keeping pace with rapid technical change. Developing practical policy instruments, external guidance for industry, and internal acquisition templates and SOPs reduces inconsistency and improves audit readiness.

Issuing Formal AI Policy Guidance to Industry

Agencies should issue formal guidance to industry clarifying expectations for acceptable AI use in proposal preparation and performance, required disclosures and metadata, security and privacy constraints for contractor tools, and how AI-related claims will be evaluated. This guidance can be delivered through acquisition policy memoranda, standardized solicitation attachments, and industry day materials. It should include concrete examples of what constitutes "material AI use," what evidence is expected for AI-produced analyses, and what logs or artifacts must be retained for post-award audit. Doing so reduces ambiguity, levels the playing field, and supports OMB's push for agency AI governance and risk management (OMB, 2024).

External organizations have also recommended shared guidance mechanisms that streamline safe AI procurement across government, such as a national guidance platform for AI acquisition (RAND, 2025). Agencies can adopt this approach by publishing a lightweight "AI Acquisition Guide for Vendors" and updating it regularly. In addition, the government could include a standard "AI Instructions to Offerors" attachment for AI-relevant procurements that covers disclosure, data handling, security, and evaluation evidence requirements (NIST, 2024).

Developing Internal Federal AI Acquisition Policies, Templates, and SOPs

Internally, agencies should build a reusable AI acquisition kit—templates, clauses, evaluation factors, checklists, and SOPs—that aligns with government and agency-wide direction and scales across contracting activities. The kit should include an "AI Oversight Annex" template for acquisition plans and source selection plans; standard language for VV&A evidence, monitoring, and audits; a QASP addendum for AI-enabled contractor workflows and deliverables; and incident response playbooks for AI failures.

OMB's M-24-10 memorandum provides a baseline for governance and risk management practices that should be reflected in internal procedures (OMB, 2024). NIST's Generative AI profile offers actionable risk categories and control ideas that can be translated into SOP checklists and acceptance criteria (NIST, 2024). GAO's accountability practices can serve as an audit-aligned backbone for these templates, ensuring traceability and continuous monitoring (GAO, 2024). GAO's recent work also highlights how data quality and workforce capability determine whether AI can be leveraged responsibly in oversight-heavy domains, lessons that apply directly to acquisition integrity and payment controls (GAO, 2026). Finally, agencies should implement a "tiered authorization" model in which staff may use low-risk AI tools by default, but acquisition use cases with material decision influence require documented approval and training authorization.

AI Literacy for the Acquisition Workforce

Human agency fails in practice when the workforce cannot recognize AI limitations, challenge outputs, or operate monitoring controls. AI literacy for acquisition is not a generic introduction to AI; it is role-based competence tied to acquisition decisions. GAO has noted that the rapid expansion of generative AI use across agencies creates management challenges and

underscores the need for policy-compliant adoption and workforce readiness (GAO, 2025). DAU reports similar findings, highlighting practical GenAI use cases in acquisition and the importance of training to use these tools effectively (DAU, 2025). HAAF recommends a role-based literacy model aligned to acquisition roles:

- **All hands (baseline):** understanding what AI can or cannot do; data sensitivity; hallucination risk; prompt hygiene; and mandatory disclosure rules.
- **Practitioners (1102s, CORs, analysts):** interpreting AI outputs; detecting red flags; validating citations; understanding uncertainty; and documenting AI use in acquisition records.
- **Supervisors and Source Selection Authority (SSA) support staff:** selecting oversight tiers; setting acceptance criteria; approving use cases; and managing risk registers.
- **Technical and governance leads (CAIO/IT/security):** model governance; monitoring; security testing; incident response; and audit readiness.

A practical literacy requirement is the ability to run "sanity checks" and structured reviews. For example, when AI drafts evaluation narratives, reviewers should confirm that every factual claim traces to the proposal or record, the narrative aligns with stated evaluation factors, and the tool does not introduce prohibited comparisons or non-evaluated criteria. When AI supports market research, reviewers should check source diversity, recency, and whether the tool's summary omits disconfirming evidence. Finally, literacy must include organizational learning loops. MITRE (2024) argues that acquisition should blend the "art" of expert judgment with AI's ability to reduce tedious tasks, provided organizations proceed cautiously with testing and integration. Institutionalizing after-action reviews, override analysis, and periodic refresher training helps sustain this cautious adoption posture.

Every agency should establish mandatory, role-based AI training tied to authorization for using Tier 1 or Tier 2 tools; incorporate AI use and oversight procedures into SOPs and Quality Assurance Surveillance Plans (QASPs); and conduct periodic exercises, such as tabletops or red teams, to practice intervention and documentation under realistic acquisition scenarios (DAU, 2025).

AI can improve acquisition speed and analytical capacity, but only if governance preserves accountable human judgment. The HAAF provides an actionable approach by defining oversight tiers for each use case, requiring VV&A evidence proportional to risk, and equipping the workforce with the knowledge to challenge and supervise AI systems. These steps align with current federal governance direction and emerging assurance guidance, while protecting the integrity and defensibility of acquisition outcomes.

Operationalizing the Framework: Use Cases

Conceptual frameworks for responsible AI adoption in federal acquisition must ultimately be validated through practical application. While the preceding sections identified risks inherent in an AI-to-AI procurement ecosystem, acquisition leaders require concrete examples demonstrating how AI can be integrated into real-world workflows without displacing human judgment or undermining accountability. This section operationalizes the proposed framework through applied use cases that illustrate how narrowly scoped AI capabilities, implemented as micro-tools, can enhance acquisition effectiveness while preserving transparency, traceability, and human agency.

Veterans Affairs Example

The Department of Veterans Affairs (VA) provides a compelling hypothetical example of how AI can be responsibly integrated into early-phase acquisition activities while strictly preserving human agency. In a recent market research effort, VA's acquisition team issued a Request for Information (RFI) to gather industry insights on a complex, mission-critical capability, supporting Veteran healthcare delivery through advanced technological solutions (e.g., interoperability, data analytics, or clinical workflow enhancements aligned with VA's digital modernization goals). The RFI elicited over 75 detailed vendor responses, each averaging 50 pages of technical narratives, capability statements, past performance examples, and innovative concepts. Manually reviewing and synthesizing this volume of unstructured, heterogeneous material would have required weeks of effort from a small team of contracting specialists, program managers, and SMEs, risking fatigue-induced oversights and schedule slippage in an environment where rapid requirements refinement is essential to accelerating Veteran capabilities.

To address this challenge without ceding evaluative control, VA deployed a narrowly scoped AI micro-tool, a generative AI-assisted analysis pipeline configured as a decision-support utility rather than an autonomous evaluator. The tool, built on secure LLM infrastructure with strict data handling protocols, performed the following bounded functions:

- **Key Data Extraction:** The AI parsed each RFI response to automatically extract structured elements such as proposed technical approaches, compliance with VA technical reference models, estimated maturity levels (e.g., TRL/MRL), teaming structures, cybersecurity postures, and rough-order-of-magnitude (ROM) cost ranges.
- **Theming and Clustering:** Using natural language processing techniques, the tool identified emergent themes across responses (e.g., common barriers to interoperability, recurring emphasis on cloud-native architectures, concerns about legacy system integration, or innovative uses of predictive analytics for Veteran outcomes). It grouped similar concepts, quantified theme prevalence (e.g., "65% of respondents highlighted FHIR-based standards"), and surfaced outliers (e.g., novel, non-traditional approaches from small businesses).
- **Summary Reporting:** The AI generated concise, evidence-based synopses for each response, with hyperlinks to original document sections, page numbers, and quotes. Importantly, the micro-tool operated under explicit HITL constraints consistent with the proposed HAAF Tier 1 oversight: AI outputs were never used directly in decision-making. All generated summaries, themes, and extractions were treated as preliminary working aids requiring mandatory human review and validation.

The acquisition team, consisting of 1102s, CORs, SMEs, and clinical/technical leads, then took ownership of the process through a structured, multi-stage workflow:

1. **Initial Validation Sprint:** In collaborative sessions, evaluators cross-checked a statistically sampled subset (approximately 20%) of AI-extracted data against source documents to confirm accuracy, detect any hallucinations or misinterpretations (e.g., conflating vendor marketing language with verifiable capabilities), and refine prompt parameters for subsequent runs. This calibration step achieved over 95% alignment after one iteration.
2. **Thematic Synthesis and Gap Analysis:** Human reviewers examined the AI-generated theme clusters to draw contextual conclusions. For instance, while the tool highlighted widespread industry support for API-first designs, VA experts identified that many responses underestimated VA-specific data sovereignty and Veteran privacy

requirements under 38 U.S.C. and HIPAA. This human insight led to strengthened requirements language around federated identity and zero-trust principles.

3. **Impact on Requirements Development:** The combined AI-assisted theming and human evaluation directly informed refinement of the draft Performance Work Statement (PWS) and evaluation criteria. Feedback revealed industry concerns about VA's legacy integration challenges, prompting adjustments to emphasize modular, open-architecture solutions and phased deployment incentives. It also surfaced opportunities for non-traditional entrants, leading to targeted small business set-aside considerations and outreach to underrepresented vendors.
4. **Documentation and Traceability:** Every final decision, whether accepting, modifying, or rejecting an AI-flagged theme, was logged with rationale, linked back to primary source material, and incorporated into the acquisition plan's AI Oversight Annex. This confirmed full defensibility during internal reviews, potential protests, or GAO scrutiny.

The outcome was transformative yet fully human-centered: market research that traditionally consumed 8–12 weeks was compressed to approximately 2–3 weeks, freeing capacity for deeper industry engagement (e.g., one-on-one vendor dialogues and solution demonstrations). More critically, the AI micro-tool acted purely as an accelerator for pattern recognition and data reduction, never as a substitute for judgment. Human experts retained authority to interpret qualitative nuances (e.g., assessing the credibility of a vendor's claimed innovation versus hype), weighed mission alignment against Veteran outcomes, and made binding determinations on requirements direction.

This VA use case exemplifies HAAF principles in practice: Tier 1 HITL for material market research influencing downstream source selection; scaled VV&A through sampling, calibration, and audit logging; and role-based AI literacy enabling the workforce to challenge outputs effectively. By treating AI as a force multiplier for human insight rather than a decision proxy, VA preserved accountability, mitigated risks of homogenization or bias amplification, and confirmed that industry feedback meaningfully shaped an acquisition approach better aligned with accelerating high-impact capabilities for Veterans. This disciplined integration demonstrates that responsible AI adoption can enhance speed and analytical depth in federal acquisition without eroding the human command essential to ethical, mission-successful outcomes.

FFRDC Example

A central reference point for the HAAF is the Acquisition AI Research (ACQAIRE) initiative developed by the MITRE Corporation. ACQAIRE provides a research and experimentation environment focused specifically on acquisition-relevant AI use cases, enabling government, industry, and FFRDC stakeholders to explore how AI can support acquisition tasks under disciplined governance and human-in-the-loop conditions.

Micro-Tools as Decision Support, Not Decision Replacement

A foundational principle of the proposed framework is that AI should be deployed as a collection of narrowly scoped, task-specific micro-tools rather than as an end-to-end decision-making system. ACQAIRE exemplifies this approach by emphasizing AI capabilities that assist acquisition professionals in performing discrete functions, such as document triage, evidence identification, or cross-volume consistency analysis, without automating evaluative judgments.

Within an ACQAIRE-like environment, AI tools are deliberately constrained to surface information rather than render conclusions. For example, instead of generating composite scores or ranking offerors, AI may identify passages relevant to specific evaluation factors or flag potential areas of interest for further human review. Evaluators retain full authority to interpret the evidence, weigh trade-offs, and determine relevance to mission needs. This design

choice directly mitigates automation bias and reinforces the role of acquisition professionals as the final arbiters of value and risk.

This micro-tool paradigm is transferable across acquisition contexts. Government acquisition offices can use such tools to manage evaluator workload during complex source selections; industry teams can employ similar capabilities internally to improve proposal coherence and compliance without overstating claims; and FFRDCs can serve as neutral venues for testing and validating tool behavior before broader adoption. In each case, AI augments cognition and efficiency without supplanting judgment.

Evidence Tagging and Traceability in Evaluation Workflows

One of the most significant contributions of ACQAIRE-aligned use cases is the emphasis on evidence tagging and traceability. A persistent challenge in AI-assisted acquisition is ensuring that outputs remain directly linked to authoritative source material. ACQAIRE addresses this challenge by associating AI-identified observations, such as potential strengths, weaknesses, or risks, with specific proposal text, page locations, and evaluation criteria.

This approach enhances transparency and legal defensibility. Evaluators can review AI-suggested annotations, accept or reject them, and document their reasoning while maintaining a clear audit trail that demonstrates how conclusions were reached. During debriefings, protests, or oversight reviews, acquisition officials can point to documented evidence chains rather than opaque model outputs. AI thus becomes a tool for strengthening documentation rigor rather than obscuring decision logic.

Evidence tagging also reduces the risk of compounding errors throughout the life cycle. By ensuring that summaries and analytical outputs remain anchored to original proposal content, ACQAIRE-style tools mitigate the effects of hallucination, over-compression, and mischaracterization. Evaluators are consistently prompted to engage with primary source material, reinforcing human oversight and contextual understanding.

Evaluator Support and Cognitive Load Reduction

Complex source selections place substantial cognitive demands on evaluation teams, particularly in time-constrained acquisition environments. ACQAIRE use cases demonstrate how AI micro-tools can reduce cognitive load while preserving evaluative rigor. Examples include tools that highlight inconsistencies across proposal volumes, flag potential disconnects between technical and management approaches, or identify deviations from solicitation requirements.

Importantly, these tools do not determine compliance or merit. Instead, they function as structured prompts that help evaluators focus their attention where it is most needed. By reducing time spent on mechanical cross-checking, AI enables acquisition professionals to concentrate on higher-order judgments involving innovation, risk, and mission alignment. This supports acquisition speed without sacrificing the qualitative assessments essential to warfighting effectiveness.

From a workforce development perspective, this approach also preserves and reinforces expertise. Junior evaluators gain insight into how experienced professionals assess evidence and trade-offs, while senior evaluators remain responsible for final judgments. AI's role is explicit and bounded, supporting learning rather than eroding skill.

Applicability Across Government, Industry, and FFRDC Environments

The ACQAIRE model demonstrates that the proposed framework is feasible across diverse acquisition environments. In government acquisition offices, AI micro-tools can be integrated into existing evaluation processes with minimal disruption, provided that governance

mechanisms for logging, version control, and human-in-the-loop decision points are established. In industry settings, similar tools can be applied during internal reviews to identify unsupported claims, inconsistencies, or compliance gaps before proposal submission, reducing downstream execution and accountability risks.

FFRDCs such as MITRE play a distinctive role in this ecosystem. Positioned at the intersection of policy, technology, and acquisition practice, they provide neutral environments for experimentation, red-teaming, and validation. ACQAIRE illustrates how FFRDC-led research efforts can translate high-level AI policy principles into operationally viable tools, offering acquisition organizations a lower-risk pathway to responsible AI adoption.

Implications for Framework Adoption

The ACQAIRE use cases demonstrate that the AI-to-AI procurement paradox is not an unavoidable consequence of automation, but rather the result of design and governance choices. When AI is implemented as a set of constrained, transparent, and traceable micro-tools, it can enhance evaluation quality and acquisition speed while preserving accountability and human agency.

By grounding the conceptual framework in real-world practice, ACQAIRE shows that responsible AI adoption in federal acquisition is achievable using existing technologies. The key is disciplined integration: clearly defined human-in-the-loop thresholds, robust evidence traceability, and deliberate limitation of AI authority, as outlined by the proposed HAAF.

Recommendations and Policy Implications

Establish a Federal AI Acquisition "Circuit Breaker" Protocol: The HAAF prescribes oversight tiers but lacks a mechanism for addressing AI system failures during production in mid-acquisition. A circuit breaker policy would require agencies to predefine automatic suspension triggers, such as a confidence score dropping below a validated threshold, a drift indicator activating, or a protest citing AI-generated content. These triggers would immediately pause AI-assisted evaluation and revert to manual processes. The key insight is that the HAAF's assurance gates are pre-deployment controls; this protocol fills the gap for real-time failure response during active source selections, where stopping and restarting can have significant cost and schedule consequences.

Establish an Independent Federal AI Acquisition Auditor Function: The HAAF's Gate D calls for independent audits but does not assign any institution the authority or resources to conduct them. An independent Federal AI Acquisition Auditor, potentially within GAO, would have standing authority to audit AI tool deployments in federal acquisition, review protest records for AI-related patterns, and publish public findings. The critical design principle is that this function must be independent of the agencies being audited, which the HAAF's current internal board structure does not guarantee. This mirrors the logic behind the Inspector General system: internal governance is necessary but not sufficient when public accountability is at stake.

Integrate HAAF Tiers into DODI 5000.02: The Under Secretary of War for Acquisition and Sustainment (USW(A&S)) should amend DOD Instruction 5000.02 to require HAAF-tier assignments across MDAPs, OTAs, and rapid acquisition pathways. This memo establishes policy and prescribes procedures for managing acquisition programs from initiation through sustainment (DODI 5000.02, 2020). While DODI 5000.02 defines milestone decision authorities, documentation requirements, and oversight thresholds for programs, it does not address how AI tools used within those acquisition processes should be governed.

Establish a Sunset and Revalidation Requirement for Deployed AI Acquisition Tools: The HAAF addresses pre-deployment validation and ongoing monitoring but lacks a mechanism for

periodic revalidation. AI models degrade, training becomes outdated, and acquisition environments evolve; a tool validated in 2024 may behave very differently when applied to 2027 proposals, potentially without any visible failure signal. A mandatory sunset provision would require every AI tool used in acquisition to undergo full revalidation against current data on a defined cycle, with automatic suspension if revalidation is not completed. This closes what is perhaps the most institutionally underappreciated gap in the HAAF: the assumption that a tool that passed assurance gates at deployment remains trustworthy indefinitely. To date, this has not been tested in a protest or court venue, and new limits on the use of AI may be imposed later depending on such outcomes.

Conclusion: Keeping Humans in Command

AI will continue to transform federal acquisition by dramatically expanding the government's ability to process information, evaluate alternatives, and accelerate procurement timelines. However, integrating AI into acquisition workflows introduces governance challenges that cannot be addressed through technology alone. As automated tools increasingly influence both proposal generation and evaluation, maintaining meaningful human agency becomes essential to preserving the integrity, transparency, and accountability of federal procurement. "The future of federal procurement is about achieving mastery in both AI efficiency and human judgment. The winners will be those organizations that act now" (Shipley Associates, 2026).

This paper introduced the concept of the AI-to-AI procurement paradox to describe the emerging interaction between government and industry AI systems within acquisition processes. If left unmanaged, this risks amplifying bias, obscuring evaluative reasoning, and eroding the expert judgment required for mission-aligned procurement decisions. To mitigate these risks, the HAAF provides a structured governance approach for integrating AI into acquisition while preserving human control. By establishing explicit oversight tiers, implementing verification and audit mechanisms, strengthening explainability and traceability requirements, embedding governance expectations into policy and contracts, and developing an AI-literate acquisition workforce, agencies can responsibly harness AI capabilities without compromising procurement integrity.

Ultimately, the objective is not to slow the adoption of AI within federal acquisition, but to guide it. The warfighter's need for speed must be balanced with institutional safeguards that confirm acquisition decisions remain defensible, ethical, and strategically sound. When implemented responsibly, AI can serve as a force multiplier for human expertise rather than a substitute for it. Keeping humans in command confirms that technological acceleration strengthens, rather than undermines, the trust and accountability upon which effective defense acquisition depends.

References

Ada Lovelace Institute. (2024, October 1). *Buying AI: Is the public sector equipped to procure technology in the public interest?* https://www.adalovelaceinstitute.org/report/buying-ai-procurement/

Bode, I., & Huelss, H. (2025). Establishing human responsibility and accountability at early stages of the lifecycle for AI-based defence systems. *Ethics and Information Technology*.

DAU. (2025). *Using AI to rapidly improve acquisition outcomes*. https://www.dau.edu/4edacm/newsletters/jan-2025/using-artificial-intelligence

DOW. (2026). *AI strategy for the Department of War*. https://media.defense.gov/2026/Jan/12/2003855671/-1/-1/0/ARTIFICIAL-INTELLIGENCE-STRATEGY-FOR-THE-DEPARTMENT-OF-WAR.PDF

European Parliament and Council. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. https://data.europa.eu/eli/reg/2024/1689/oj

FNN. (2025). *AI to the rescue of federal procurement?* https://federalnewsnetwork.com/federal-insights/2025/09/ai-to-the-rescue-of-federal-procurement

GAO. (2024). *Artificial intelligence: AI accountability framework and related resources*. https://www.gao.gov/artificial-intelligence

GAO. (2025). *Generative AI use and management at federal agencies*. https://www.gao.gov/products/gao-25-107653

GAO. (2025). *Artificial intelligence: Federal efforts guided by requirements and advisory groups* (GAO-25-107933). https://www.gao.gov/assets/gao-25-107933.pdf

GAO. (2026). *Fraud and improper payments: Data quality and a skilled workforce are key to leveraging AI*. https://www.gao.gov/products/gao-26-108850

GSA. (n.d.). *Buy AI — GSA OneGov AI acquisition resources and contracting vehicles*. https://www.gsa.gov/technology/government-it-initiatives/artificial-intelligence/buy-ai

GSA. (2025). *Use of artificial intelligence at GSA* (CIO Order No. CIO 2185.2). https://www.gsa.gov/directives-library/use-of-artificial-intelligence-at-gsa

Hubbard, R. (2025). *How AI can accelerate defense acquisition and help U.S. forces keep pace with technology*. Federal News Network.

iQuasar. (2026). *AI procurement in 2026: How federal agencies will use automation to evaluate and monitor contracts*. https://iquasar.com/blog/ai-procurement-in-2026-how-federal-agencies-will-use-automation-to-evaluate-and-monitor-contracts

Johnson, J. (2026). Can AI behave ethically during military crises? Preserving human moral agency. *International Affairs*.

Kolt, N. (2025). *Lessons from complex systems science for AI governance*. PMC - NIH. https://pmc.ncbi.nlm.nih.gov/articles/PMC12365527

MeriTalk. (2026). *GSA rethinks federal acquisitions as AI reshapes contracting*. https://www.meritalk.com/articles/gsa-rethinks-federal-acquisitions-as-ai-reshapes-contracting

MITRE (2024). *Enhancing acquisition outcomes through leveraging of artificial intelligence*. https://www.mitre.org/sites/default/files/2025-03/PR-24-0962-Leveraging-AI-Acquisition.pdf

NIGP. (2026, February 11). *When artificial intelligence buys for us — Transparency and accountability in AI-driven procurement*. https://www.nigp.org/blog/when-ai-buys-for-usv

National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST AI 100-1). https://doi.org/10.6028/NIST.AI.100-1

National Institute of Standards and Technology. (2024). *Artificial intelligence risk management framework: Generative artificial intelligence profile* (NIST AI 600-1). https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf

Office of Management and Budget. (2024, March 28). *Advancing governance, innovation, and risk management for agency use of artificial intelligence* (M-24-10). Executive Office of the President. https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf

Office of the Under Secretary of Defense for Acquisition and Sustainment. (2020, January 23). *DoD Instruction 5000.02 Operation of the Adaptive Acquisition Framework*. Department of Defense. https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500002p.PDF

Open Contracting Partnership. (2025, November 10). *Governments risk "expensive failures" without smarter AI procurement, new report warns*. https://www.open-contracting.org/news/governments-risk-expensive-failures-without-smarter-ai-procurement-new-report

Shipley Associates. (2026). *The human touch returns: How thoughtful AI integration is reshaping federal procurement*. https://www.shipleywins.com/blogs/the-human-touch-returns-how-thoughtful-ai-integration-is-reshaping-federal-procurement

The White House. (2025). *White House unveils America's AI Action Plan.* https://www.whitehouse.gov/articles/2025/07/white-house-unveils-americas-ai-action-plan/

Tillipman, J. (2026). Abdicated judgment: AI tools and the future of reasoned decision-making in federal procurement. *Yale Journal on Regulation Notice & Comment*. https://www.yalejreg.com/nc/abdicated-judgment-ai-tools-and-the-future-of-reasoned-decision-making-in-federal-procurement-by-jessica-tillipman

U.S. DoS. (2024). *Department of State compliance plan for OMB Memorandum M-24-10*. https://2021-2025.state.gov/wp-content/uploads/2024/09/DOS-Compliance-Plan-with-OMB-M-24-10-Accessible-9.23.2024.pdf


Acquisition Research Program
Department of Acquisition, Finance, and Manpower
Naval Postgraduate School
555 Dyer Road, Ingersoll Hall
Monterey, CA 93943

www.acquisitionresearch.net