

Artificial Intelligence in Defense

Data Integrity, Operational Utility, and Strategic Risk

LT COL RYAN BLAKENEY, USAF

Abstract

The Department of War is integrating artificial intelligence (AI) at a pace that outstrips institutional capacity to validate the technology and develop the force to wield it safely. The link between AI capability and operational employment is not technical expertise, but data literacy. Operators and commanders must understand the origin and fidelity of information underpinning their models. Without this literacy, recognizing when foundational data has been corrupted or injected by an adversary becomes exceptionally difficult. Using the Air Force's wargaming enterprise and the Collaborative Combat Aircraft program as case studies, this article exposes three interlocking vulnerabilities: data integrity concerns, a trust gap as simulations transition to physical platforms, and institutional shortfalls in professional military education. To bridge this gap, this article proposes a five-tier Acceptable Level of Risk schema, adapted from Air Force risk management doctrine, equipping commanders with a shared vernacular to assess, communicate, and manage AI integration risk.

Keywords: artificial intelligence (AI) integration, data integrity, data literacy, operational risk, Acceptable Level of Risk (ALR), adversarial data manipulation, wargaming and simulation, autonomous systems / collaborative combat aircraft (CCA), human-machine trust, professional military education (PME)

Introduction

The Department of War (DOW) stands at a strategic inflection point. The adoption of artificial intelligence (AI) and large language models has crossed from experimentation into operational reality, marking the early contours of a new military offset. AI no longer resides primarily in the laboratory. It now shapes planning, intelligence synthesis, simulation, and increasingly, the conduct of operations themselves. Autonomous mission software is moving steadily into Collaborative Combat Aircraft (CCA) programs, while edge processing of sensor data has become an operational necessity in contested and degraded environments. In such conditions, adversaries deny command-and-control networks and tactical data links, making reliance on reach-back both impractical and dangerous. At the same time, large language models increasingly support wargame adjudication and operational planning, compressing analytic timelines and offering decision aids once dependent on large staffs and extended deliberation. The speed of fielding has been striking. The central challenge, however, lies not in the technology but in the widening gap between what these systems can produce and what commanders and operators can meaningfully verify under operational time constraints.

Operators who cannot trace an AI system's outputs to their source data, assess the conditions under which the model was trained, or recognize when it operates beyond its validated parameters do not act as safeguards against error. They become conduits for it. In a commercial setting, the consequences amount to little more than a poor recommendation or an irrelevant advertisement. In war, the cost is far higher. It may take the form of a corrupted targeting solution, a wargame that generates strategically unsound plans with unwarranted mathematical confidence, or an autonomous platform executing engagements informed by poisoned or manipulated data. These failures do not register on quarterly earnings statements. They register in lost lives, failed missions, and unfavorable strategic outcomes.

To confront this problem, this article examines three interlocking dimensions of the AI integration gap. The first is data integrity. This dimension addresses the foundational question of whether the data underpinning AI systems is accurate, current, and protected against corruption, whether through deliberate adversarial manipulation or unintentional misuse by inadequately trained personnel. The second dimension is operational utility. Here the challenge lies in extending trust from the controlled, virtual environment of wargaming into the physical domain of unmanned aerial systems and CCAs, where error carries immediate and irreversible consequences. The third dimension is institutional. Professional military education (PME) curricula and assignment policies have not kept pace with technological change and have failed to produce operators with the data literacy, discipline, and skepticism required to employ AI responsibly.

The article concludes by proposing a five-tier ALR framework for AI integration, adapted from existing Air Force risk management doctrine. This framework does not function as a bureaucratic checklist. It serves instead as a shared vocabulary that allows commanders to identify, communicate, and consciously accept the risk inherent in employing AI systems under conditions that often exceed the verifiable limits of human understanding. The framework is especially relevant in combat environments defined by time pressure, where commanders cannot fully validate the underlying data and must decide, in moments that carry life-and-death consequences, whether to trust the output provided by a machine.

The Data Foundation and Strategic Risk

AI derives its power not from code alone, but from the data on which it is built. In military applications, that dependency transforms data from a technical input into a strategic vulnerability. The integrity, provenance, and stewardship of information now shape the reliability of AI-enabled decisions long before any algorithm produces an output. Weaknesses in this foundation do not remain confined to laboratories or acquisition pipelines. They propagate outward, influencing opera-

tional trust, shaping institutional behavior, and creating openings that adversaries can exploit at scale. Understanding how data quality, adversarial manipulation, and organizational incentives interact is therefore essential to assessing the strategic risk posed by AI in war.

The Quality Problem

Every AI system rests on a single foundation: its training data. In the commercial sector, large language models and machine-learning systems draw strength from vast quantities of freely available or commercially acquired information. The military operates under no such conditions. Operational data remains fragmented across organizations, stored in incompatible formats, and classified at multiple levels within compartmented environments. The Department of Defense recognized this constraint in its 2023 *Data, Analytics, and Artificial Intelligence Adoption Strategy*, which acknowledged that these structural realities complicate the development of cohesive AI models and impede the reliable sharing of information across classification boundaries.¹

This problem is not merely technical. It is institutional. In December 2023, the Government Accountability Office (GAO) reported that the Department of Defense could not clearly identify its AI workforce, could not forecast future workforce requirements, and continued to store critical operational data in siloed systems that inhibit information sharing.² The department has also failed to conduct a comprehensive assessment of the risks associated with relying on external providers for foundational AI capabilities. The 2023 adoption strategy states plainly that high-quality, decision-ready data is a prerequisite for reliable AI deployment. Yet the institutional mechanisms required to produce, curate, and protect such data remain insufficiently developed.³

The volume illusion further compounds this challenge. Within the defense acquisition system, a persistent assumption holds that larger datasets inevitably produce better AI. That assumption is false. The relationship between the quantity and quality of training data is not linear. A vast dataset of mislabeled or poorly curated imagery will produce a model that operates with high confidence yet delivers consistently wrong results. A smaller dataset, carefully curated and validated against operational reality, will often yield a far more reliable and militarily useful system.

¹ DoD *Data, Analytics, and Artificial Intelligence Adoption Strategy* (Washington: Department of Defense, November 2023), 4, <https://media.defense.gov/>.

² *Artificial Intelligence: Actions Needed to Improve DOD's Workforce Management*, GAO-24-105645 (Washington: GAO, December 2023), 1, <https://www.gao.gov/>.

³ DoD *Data, Analytics, and Artificial Intelligence Adoption Strategy*, 4.

Quality remains decisive, even though scale retains importance when sensors encounter conditions or environments that fall outside their original training parameters.

This distinction carries particular importance for intelligence, surveillance, and reconnaissance (ISR) applications. In these contexts, an AI-enabled edge processing system may generate a targeting recommendation without any intermediate human review of the underlying data quality. Such architectures introduce significant operational risk. A system trained over a decade to identify a particular structure as a valid military target does not possess the contextual awareness required to recognize that the same structure may later be repurposed as a hospital. Absent human validation, the model will confidently recommend a strike against a protected civilian site, relying entirely on historical training data divorced from present reality. Precision in execution offers no protection when the premise itself is wrong.

Adversarial Data Manipulation

Internal data management failures represent only part of the risk. A more serious strategic threat arises when adversaries deliberately manipulate the data used to train AI systems. When an attacker alters information inside the training pipeline, that corruption does not remain transient. The model absorbs it, reinforces it, and embeds it permanently within the logic used to generate operational decisions. Errors of this kind do not degrade gradually. They persist silently until the system encounters the precise conditions under which it fails.

One of the most dangerous forms of this manipulation is data poisoning, in which an adversary subtly alters training inputs before a model is deployed. Researchers at the Georgetown Center for Security and Emerging Technology have demonstrated that imperceptible alterations to imagery can cause neural networks to misclassify targets with high consistency.⁴ These demonstrations occurred under controlled research conditions and have not been publicly confirmed in combat environments, in part because successful attacks of this nature would be difficult to detect and unlikely to be publicly attributed.⁵ The underlying vulnerability nonetheless remains real. An adversary that modifies the signatures of armored vehicles during data collection causes the model to learn an incorrect baseline, rendering it effectively blind to those systems in combat. If introduced into the

⁴ Andrew J. Lohn, *Poison in the Well: Securing the Shared Resources of Machine Learning* (Washington: Georgetown University Center for Security and Emerging Technology, June 2021), <https://cset.georgetown.edu/>.

⁵ Li Ang Zhang et al., *Operational Feasibility of Adversarial Attacks Against Artificial Intelligence*, RRA8661 (Santa Monica, CA: RAND Corporation, 2022), 1, <https://www.rand.org/>; and “SABER: Securing Artificial Intelligence for Battlefield Effective Robustness” (program description, US Defense Advanced Research Projects Agency, n.d.), <https://www.darpa.mil/>.

training pipeline of an ISR platform, such manipulation could lead the onboard AI to misidentify a surface-to-air missile site as benign infrastructure or to classify a protected civilian location as a legitimate military target, with potentially catastrophic consequences should a strike be executed.

A second and equally dangerous vector arises through evasion attacks. In this case, adversaries target AI systems during real-time operations rather than during training. By presenting sensors with carefully crafted or manipulated inputs, an opponent can induce an autonomous system to detect threats that do not exist or to overlook genuine threats present within its field of view. These vulnerabilities extend beyond theoretical concern. They constitute a new avenue for strategic attack, one that exploits the same mathematical properties that enable AI to function at operational speed and scale. The DOW's *Artificial Intelligence Cybersecurity Risk Management Tailoring Guide* formally acknowledges this danger, warning that adversarial queries can manipulate an AI system's own monitoring and defensive mechanisms against itself.⁶

Congress has also acknowledged the severity of this risk. Section 1512 of the National Defense Authorization Act for Fiscal Year 2026 directed the DOW to develop a comprehensive policy governing AI and machine-learning cybersecurity within a defined timeline. The legislation explicitly identified data poisoning, adversarial attacks, model tampering, model jailbreaks, and adversarial prompt injection as threat categories requiring formal mitigation.⁷ Legislative recognition alone does not reduce vulnerability. It does, however, establish data integrity as a documented and consequential weakness, one serious enough to demand an institutional response at the departmental level.

The Grassroots–Contractor Gap

Some of the most operationally relevant AI development occurs at the lowest levels of the force, driven by those who employ the systems daily and remain closest to the data environments those systems must navigate. In many cases, operators have begun building locally secured large language model environments on classified networks, relying on self-directed learning or attempting to adapt commercial tools to military purposes. The underlying risk is not innovation itself but opacity. The military often lacks visibility into how commercial models were trained, what data they ingested, or whether that data was ever assessed for adversarial

⁶ DoD *AI Cybersecurity Risk Management Tailoring Guide* (Washington: Chief Digital and Artificial Intelligence Office, July 2025), <https://dodcio.defense.gov/>.

⁷ National Defense Authorization Act for Fiscal Year 2026, Pub. L. No. 119-60, § 1512 (18 December 2025), <https://www.congress.gov/>.

contamination. Commercial developers rarely treat data poisoning as a central design concern because their principal risk is financial loss, not the prospect of lethal error or the misidentification of protected civilian sites. When pushed to the tactical edge without rigorous data literacy, such systems introduce unacceptable operational danger.

To mitigate this risk, bottom-up efforts by junior personnel have sought to employ AI in controlled and disciplined ways. These operators cultivate an instinctive understanding of what constitutes an operationally illogical output, an intuition that no contractor working in a sanitized laboratory can easily replicate. Years of experience at the tactical edge allow them to recognize immediately when a model's output diverges from physical reality. Their value lies not in technical sophistication alone, but in the ability to reconcile machine-generated conclusions with lived operational truth.

Despite their operational value, these grassroots capabilities face persistent structural resistance. They compete for funding, authority, and visibility against large, centralized acquisition programs designed to reward scale and formal process rather than tactical relevance. Defense contractors operate within a different incentive structure. They can optimize models against abstract performance metrics in controlled environments that lack the friction, ambiguity, and degradation inherent to combat operations. Corporate investment follows financial return and contractual milestones, not survivability under fire. Grassroots military innovation, by contrast, emerges from the demand for systems that work when conditions deteriorate and consequences turn lethal. The Defense Innovation Board has repeatedly identified this tension between bottom-up operational ingenuity and top-down acquisition processes as a recurring problem within DOW software development.⁸

The result is a predictable outcome. The department fields AI systems that perform exceptionally well under the conditions for which they were designed, yet degrade rapidly in the environments they are most likely to encounter. This gap reflects a familiar truth of warfare: the enemy adapts. Commercial systems rarely account for an intelligent, hostile adversary intent on deception, disruption, and destruction. In the commercial domain, competition aims to capture market share. In war, the adversary seeks to kill the operator and defeat the system. AI that cannot withstand that reality does not merely underperform. It becomes a liability.

⁸ Defense Innovation Board, *Software Is Never Done: Refactoring the Acquisition Code for Competitive Advantage* (Washington: Department of Defense, 3 May 2019), <https://media.defense.gov/>.

Operational Utility, From the Virtual to the Physical

AI reveals its operational value not in isolation but in use, first in simulation and ultimately in combat. As AI-enabled systems migrate from virtual environments into physical platforms, questions of trust, validation, and risk intensify. Wargaming offers a controlled setting in which to explore these tensions, allowing planners to test assumptions and evaluate machine-generated insights before consequences become irreversible. Yet the constraints of simulation also mask the very frictions that define war. When AI transitions from adjudicating abstract scenarios to executing decisions aboard aircraft and sensor platforms, errors no longer emerge as flawed plans but as physical actions. Understanding how trust is established, maintained, or misplaced across this transition is essential to assessing the operational utility of AI in modern warfare.

The Wargaming Laboratory

Wargaming has long served as the intellectual proving ground of armed conflict. It provides a contested synthetic environment in which commanders test plans, surface assumptions, and develop judgment at a cost measured in lessons learned rather than lives lost. In recent years, the demands placed on this enterprise have expanded dramatically, while its fundamental methods have struggled to evolve at the same pace.

A 2024 RAND assessment prepared for the Department of the Air Force identified a persistent structural challenge. Translating mathematical or computational representations of strategic interaction into militarily relevant problems remains a fundamental limitation of AI in wargaming.⁹ This constraint reflects the enduring tension between quantitative models and the qualitative realities of war. No model can replicate the judgment forged by decades of operational experience, account for the pilot who disregards orders because no one gets left behind, or anticipate the individual act of initiative that falls outside established doctrine and alters the course of battle. The same RAND analysis observed that AI systems often display tactical proficiency while remaining strategically naïve. The very characteristics that make games tractable for AI are those that separate them from war itself. Incomplete information, adaptive adversaries, shifting political constraints, and acts of human courage resist formalization and defy reliable quantification.

⁹ Edward Geist, Aaron B. Frank, and Lance Menthe, *Understanding the Limits of Artificial Intelligence for Warfighters*, vol. 4, *Wargames*, RRA17224 (Santa Monica, CA: RAND Corporation, 2024), vi, <https://www.rand.org/>.

Despite these structural limitations, institutional momentum behind AI-assisted wargaming continues to accelerate. The Johns Hopkins University Applied Physics Laboratory's GenWar initiative employs large language models to automate scenario generation and replay, compressing planning cycles that once spanned months into days.¹⁰ In a separate effort, the Army War College demonstrated in 2025 that AI-augmented wargaming enabled planners to synchronize forces earlier, identify risk sooner, and explore branches and sequels in greater depth than traditional methods allowed.¹¹ In both cases, however, the decisive moment of human involvement occurs at adjudication. A subject-matter expert must review AI-generated outputs before planners incorporate them into operational designs. At this juncture, the verifiability problem either reveals itself or becomes embedded in the plan. A validator must understand not only the planning process but also the provenance and limitations of the data informing the model. Without that understanding, an official may retain the authority to approve an output while lacking the competence required to audit it.

This institutional push relies heavily on multi-agent reinforcement learning systems, in which agents learn by acting within a shared environment and iterating through extensive training cycles. Frazer-Nash Consultancy illustrated the technical potential of this approach when a reinforcement learning-based red force agent achieved a 91-percent win probability against a nominally superior blue force.¹² Such outcomes demonstrate computational capability, not operational utility. As Garry Brewer and Martin Shubik cautioned in their critique of applying mathematics to warfare, the ability to solve a model offers no assurance that the model itself captures the right problem.¹³

When initial problem formulation is flawed, AI amplifies error rather than correcting it. Optimization drives the simulation toward the wrong answer with increasing confidence. One AI-enabled wargame conducted in 2020 illustrates this risk. The system failed because its training data did not reflect a critical operational constraint: the adversary's legacy computer infrastructure lay beyond the reach of

¹⁰ "Generative AI Wargaming Promises to Accelerate Mission Analysis" (news release, Johns Hopkins University Applied Physics Laboratory, 3 March 2025), <https://www.jhuapl.edu/>.

¹¹ Tom Spahr, Layton Mandle, and Mike Stinchfield, "Back to the Basics in Wargaming: With a Little Help from AI," *War Room*, 11 September 2025, <https://warroom.armywarcollege.edu/>.

¹² Geist, Frank, and Menche, *Understanding the Limits of Artificial Intelligence for Warfighters*, 18, citing FrazerNash Consultancy, "Using AI in Wargaming Simulation as a Decision Support Tool," STOMPMSG217, <https://www.fnc.co.uk/>. The 91-percent win probability derives from the FrazerNash proof-of-concept Red Force Response tool trained across 17,587 reinforcement learning episodes.

¹³ Garry D. Brewer and Martin Shubik, *The War Game: A Critique of Military Problem Solving* (Cambridge, MA: Harvard University Press, 1979), 14, cited in Geist, Frank, and Menche, *Understanding the Limits of Artificial Intelligence for Warfighters*, 12.

modern cyber capabilities.¹⁴ The model could not account for a reality absent from its inputs. As a result, it generated outputs that appeared analytically sound but rested on false premises, leading planners toward inaccurate assumptions and dangerously distorted risk assessments.

This failure reflects a deeper abstraction problem inherent to AI. AI systems operate strictly within the parameters they are given. They cannot generate genuinely novel insights or account for human decisions that fall outside those bounds. Such departures are not anomalies in warfare. They define it. War turns on judgment, improvisation, defiance of expectation, and decisions that resist formal modeling. An AI system unable to represent these features will produce plans that appear mathematically optimized while remaining operationally incomplete. The Center for Strategic and International Studies has identified this gap as a core educational challenge. As AI-assisted planning becomes more prevalent, professional military education (PME) must equip planners with the tools to interrogate data, question assumptions, and recognize when machine-generated outputs are internally coherent yet strategically wrong.

From Simulation to Physical Platform

The abstraction problem that limits artificial intelligence in wargaming intensifies when AI moves from simulation into physical systems. Modern ISR platforms collect immense volumes of sensor data. In a contested, multi-domain environment, transmitting every increment of raw information back to a combined air operations center creates a severe vulnerability. Adversaries can deny or degrade communications, rendering reach-back unreliable or impossible. AI offers a partial solution through edge processing, allowing platforms to analyze sensor data locally, identify priority information, and transmit only actionable intelligence to human decision-makers, whether positioned at the tactical edge or operating through intermittent command-and-control links.¹⁵

This same operational logic underpins the Air Force's CCA program. For these systems to function as effective sensor nodes, electronic warfare platforms, and weapon carriers in degraded environments, their onboard AI must execute autonomous flight, prioritize threats, and maneuver tactically without continuous direction from a manned aircraft. In February 2026, the Air Force contracted Collins Aerospace and Shield AI to develop mission autonomy software using an

¹⁴ Peter Byrne, "The Problem with AI War Games," *Project Censored*, 26 June 2025, <https://www.projectcensored.org/>.

¹⁵ *Final Report* (Washington: National Security Commission on Artificial Intelligence, March 2021), 84, <https://www.dwt.com/>.

architecture that separates autonomy from specific hardware, enabling rapid algorithm updates as threats evolve. The technical trajectory is clear. Autonomy is no longer constrained by engineering feasibility.

On 23 February 2026, the YFQ-44A completed its first captive-carry flight with an inert AIM-120 AMRAAM air-to-air missile. In its public release, the Air Force emphasized that a human retains authority over weapons release throughout development and testing. That assertion reflects existing policy and aligns with established risk management doctrine. It does not resolve the more difficult question of trust. A manned wingman must determine whether a targeting solution generated by an autonomous aircraft rests on valid and unmanipulated sensor data. Analysts at the Center for Strategic and International Studies have documented persistent confusion surrounding current autonomous weapons policy, noting that senior leaders have repeatedly mischaracterized the human-in-the-loop requirements of Department of Defense Directive 3000.09. They also note that the transition from semi-autonomous to fully autonomous operation could occur through a relatively minor software modification. At this stage, technology no longer constitutes the primary constraint. Decisions about autonomy now hinge on risk tolerance, policy judgment, and ethical responsibility.

The ultimate constraint on autonomy is not technological maturity but institutional preparedness. The decisive question is whether the organization has developed the frameworks, training systems, and personnel with sufficient data literacy, risk comprehension, and risk acceptance to recognize when critical transitions occur and to understand their operational consequences. Recent demonstrations underscore the immediacy of this challenge. The DARPA Air Combat Evolution program showed AI agents conducting within-visual-range air-to-air combat during multiple test flights against human pilots. Australia's MQ-28 Ghost Bat has likewise demonstrated autonomous air-to-air weapon employment as part of a loyal wingman construct.¹⁶ These developments do not describe a speculative future. They reflect capabilities already in hand.

The cognitive discipline required to validate sensor data aboard an ISR platform is the same discipline required to assess a targeting recommendation generated by a CCA. In both cases, the operator must determine the origin of the data, the conditions under which it was collected, the signatures that corrupted or manipulated inputs might exhibit, and whether the system has encountered conditions beyond its training envelope. That assessment also requires the ability to recognize adversary use of camouflage, concealment, and deception intended to manipulate

¹⁶ "Boeing, RAAF Achieve CCA Missile Fire from MQ-28 Ghost Bat" (press release, Boeing Defence Australia, 9 December 2025), <https://boeing.mediaroom.com/>.

sensors or exploit computational vulnerabilities. This form of judgment does not emerge organically from exposure to technology. The institution must teach it deliberately, reinforce it systematically, and treat it as a core operational competency rather than a technical adjunct.

The Institutional Gap

The challenges associated with AI integration ultimately converge not at the level of technology, but at the level of institutions. Systems can be fielded faster than organizations can adapt to employ them, and capability can outpace the human structures required to exercise discipline over its use. The most consequential risks examined in the preceding sections arise not because AI performs beyond specification, but because the force lacks the personnel frameworks, educational models, and assignment structures necessary to validate and govern its outputs. This institutional lag transforms technical vulnerability into strategic exposure. Without deliberate investment in workforce identification, education, and professional development, the military risks fielding increasingly autonomous systems while simultaneously eroding the very human judgment required to control them.

The Workforce Identification Problem

The problem described in the preceding sections is not primarily technological. The technology exists, continues to advance rapidly, and will be fielded regardless of institutional readiness. The central challenge is human capital. The US military fields some of the most advanced AI systems ever developed, yet it lacks the organizational infrastructure required to produce operators, planners, and commanders capable of employing those systems with discipline, skepticism, and data literacy. Before an institution can educate such a workforce, it must first be able to identify it. At present, the DOW cannot do so with confidence.

In December 2023, the GAO reported that the department had defined artificial intelligence work roles and identified certain occupations engaged in AI-related activities. It had not, however, assigned responsibility to the organizations required to complete the additional steps necessary to formalize that workforce. These steps included coding AI-related roles within personnel systems, establishing qualification standards, and updating workforce guidance. As of late 2025, the GAO continued to list these foundational recommendations as either open or only partially addressed, indicating limited institutional progress.¹⁷

¹⁷ *Artificial Intelligence: Actions Needed to Improve DOD's Workforce Management*, 1.

While the DOW acknowledged these shortfalls in an August 2024 corrective action plan but has struggled to overcome bureaucratic inertia. The Chief Digital and Artificial Intelligence Office launched an initiative intended to code 50 percent of the Total Force across personnel and manpower systems during calendar year 2025 to establish a common baseline of digital capability.¹⁸ Despite this effort, the GAO has indicated that the necessary structural fixes will not be fully implemented until at least late 2026, and possibly as late as early 2027. The department has further noted that it requires additional funding to complete these reforms. At the same time, its strategic documentation continues to rely on inconsistent terminology, such as references to an “innovation workforce,” instead of adopting standardized and enforceable AI definitions as recommended by the GAO.¹⁹

The National Security Commission on Artificial Intelligence identified this workforce gap with urgency. In its 2021 *Final Report*, the commission concluded that the shortage of artificial intelligence talent represented one of the primary obstacles to making the United States AI-ready by its projected target of 2025.²⁰ That date has now passed. The gap remains. The commission did not attribute this shortfall to a lack of technically trained personnel in the civilian sector. It identified the constraint instead as the military’s inability to attract, develop, and employ such expertise within the rigid structures of a uniformed service. Until the department resolves this basic problem of workforce identification, investments in AI risk outpacing the human capacity required to govern their use.

The Assignment Policy Gap

The institution’s human capital shortfall becomes most visible in the way the US Air Force manages personnel with advanced academic training in science and AI. Officers who complete specialized education, Joint planning programs, or rigorous academic training often return to the operational force without a designated billet structure prepared to employ them. The assignment system was not built to accommodate individually acquired data science expertise. While the Air Force recognizes certain analytically focused career fields, particularly within the operations research analyst designation and related enlisted specialties, this narrow framework fails to account for operators who combine technical training with experience at the tactical edge. The assignment architecture has not been meaningfully redesigned to recognize or employ such hybrid expertise.

¹⁸ “DoW Digital Workforce,” Chief Digital and Artificial Intelligence Office, n.d., <https://www.ai.mil/>.

¹⁹ *Artificial Intelligence: Actions Needed to Improve DOD’s Workforce Management*, 18–19.

²⁰ *Final Report*, 19.

The consequences of this misalignment are predictable. A rated officer who completes a graduate degree in data science frequently returns to a flying unit, a mandatory staff position, or another PME assignment unrelated to their new skills. Over time, technical proficiency erodes or remains unused. At the same moment, units field artificial intelligence-enabled platforms without any personnel equipped to audit system outputs, challenge the assumptions embedded in training data, or recognize early signs of model degradation or distributional shift. Capability advances, while competence lags.

The institution already possesses both the expertise and the operational demand required to close this gap. What it lacks is a mechanism to align them systematically. At present, the matching of technical skill to operational need occurs sporadically through informal networks or individual initiative rather than deliberate personnel policy. This problem extends beyond the Air Force and reflects the same structural weaknesses identified in workforce definition and education. Until the institution deliberately links specialized technical education to operational assignments, its investment in human capital will continue to generate isolated pockets of individual competence that dissipate over time, rather than building an enduring organizational capacity that compounds with experience and scale.

The PME Curriculum Gap

PME has begun to incorporate AI into its curricula, but the character of that instruction remains decisive. Familiarity with what AI can do does not equate to competence in evaluating what it is actually doing within a specific operational context. A 2025 article published by Air University argued that cultivating a data-literate force requires personnel to understand multiple levels of analytics and to apply data appropriately in terms of timing, purpose, and interpretation.²¹ While this framework effectively articulates the desired end state, it does not address the more demanding cognitive task of validating AI-generated outputs against operational ground truth.

An officer who understands the existence of predictive analytics is not necessarily prepared to judge whether a targeting recommendation derived from corrupted or misapplied training data is plausible. Analysis published by the Modern War Institute highlighted this tension, observing that data literacy must permeate leadership, logistics, and operational planning rather than remain confined to

²¹ Nathaniel W. Flack, "Data Literacy for All: Why Every Airman and Guardian Needs to Be Data Literate," *Wild Blue Yonder* (blog), 11 July 2025, <https://www.airuniversity.af.edu/>.

specialized technical instruction.²² PME programs operate under significant constraints imposed by limited contact hours, resident course requirements, and Joint PME certification mandates. Integrating meaningful instruction in data validation therefore requires deliberate curricular design and faculty with sufficient technical depth to teach the material with authority rather than superficial familiarity.

The failure to provide this education is systemic. As Iain Cruickshank argued in *Joint Force Quarterly*, the military lacks a unifying model for developing an AI-ready workforce, and existing proposals fail to account for the scale and diversity of roles that artificial intelligence integration will impose across the force.²³ That gap remains unresolved even as the DOW continues to commit substantial institutional capital to expanding AI capability. In July 2025, the Chief Digital and Artificial Intelligence Office awarded contracts totaling USD 800 million to Anthropic, Google, OpenAI, and xAI to develop agentic AI workflows across defense mission areas.²⁴ Investment in technical capability is accelerating in the absence of a corresponding investment in workforce preparation.

Some elements of the DOW have moved further than others in acknowledging this imbalance at the curricular level. In 2023, the Army Logistics University proposed a deliberate, three-tier progression in data literacy that integrates foundational instruction at the officer basic course and intermediate-level content at the captain's career course. The stated intent was to ensure planners could interrogate data rather than merely receive it. The uneven adoption of this model across the department underscores the persistence of the institutional gap. The force continues to field increasingly sophisticated AI systems without establishing an equally deliberate framework to educate those responsible for validating and employing them.²⁵

The Practitioner–Instructor Gap

A secondary problem compounds the curriculum gap. Faculty capable of teaching AI data validation at an operationally credible level remain exceptionally scarce. Developing the required depth demands either advanced education in a technical discipline or sustained, hands-on experience with the specific platforms and data environments used in Air Force operations. Most PME instructors bring substan-

²² Charlie Phelps, "Leadership, Lethality, and Data Literacy: Three Keys to Prepare the Army for the DataDriven, AIEnabled Future of War," *Modern War Institute*, 2025, <https://mwi.westpoint.edu/>.

²³ Iain Cruickshank, "An AIReady Military Workforce," *Joint Force Quarterly* 110, no. 3 (2023): 46–53, <https://ndupress.ndu.edu/>.

²⁴ Sydney J. Freedberg Jr., "Anthropic, Google and xAI Win \$200 Million Each from Pentagon AI Chief for 'Agentic AI,'" *Breaking Defense*, July 14, 2025, <https://breakingdefense.com/>.

²⁵ Bob Spivey et al., "The Future of Data Education within Army Sustainment," *Army Sustainment*, 23 February 2023, <https://www.army.mil/>.

tial operational credibility to the classroom but lack the technical background necessary to interrogate AI systems rigorously. Conversely, officers with advanced technical training rarely receive PME teaching assignments unless they participate in short-term fellowship programs that typically last no more than a year. The institution has therefore failed to establish a durable pipeline that produces instructors fluent in both operational judgment and technical validation.

This gap carries real risk. Superficial exposure to AI at the PME level can prove more harmful than no exposure at all. An officer who completes a brief instructional block and concludes that they understand AI may carry unwarranted confidence into the cockpit, the operations center, or the planning cell. That confidence becomes dangerous when planners rely on large language models to synthesize vast quantities of information without appreciating the limits of those systems. In such cases, quantitative coherence can mask qualitative error. Critical asymmetries, contextual subtleties, and design considerations slip through the cracks as planners privilege machine-generated output over the human art of linking disparate operational elements into a coherent whole.

The goal of PME should extend beyond basic familiarity with AI. It must cultivate calibrated skepticism, the capacity to incorporate AI-generated outputs as useful inputs while retaining the intellectual independence to recognize when those outputs cannot be trusted. Data literacy underpins this competency. It enables planners to understand not only how information is produced, but also why certain judgments cannot be delegated to systems that lack an appreciation for context, intuition, and the qualitative art required for sound strategic decision-making. AI can accelerate analysis, but it cannot replace judgment. PME must therefore prepare officers to resist the temptation to confuse computational confidence with operational truth.

The consequences of failing to instill this discipline extend well beyond institutional inefficiency. Adversarial attack vectors such as data poisoning, sensor manipulation, and adversarial prompting are designed precisely to generate outputs that appear plausible to a generalist while evading detection by those without technical depth. A force that fields increasingly autonomous systems while systematically underpreparing the personnel responsible for auditing their outputs creates an asymmetric advantage for any adversary willing to invest in exploiting those weaknesses. In this context, the PME shortfall is not an educational inconvenience. It is a mechanism through which a technically inferior adversary can negate technological superiority, not by matching capability for capability, but by undermining the judgment of those tasked with employing it.

Operationalizing Data Literacy: A Three-Tiered Model

Closing the institutional gap identified in the preceding sections requires more than exhortation. The military must translate data literacy from an abstract aspiration into an operational requirement. Doing so demands a deliberate progression embedded across all levels of enlisted and officer PME, with clearly defined expectations that mature alongside responsibility and authority.

The foundational tier must begin at the earliest stages of professional development, including Airman Leadership School, the Noncommissioned Officer Academy, and Squadron Officer School. At this level, AI should not appear as an isolated technical subject. It should be integrated directly into existing research and analytic methodologies so students gain practical familiarity with both the utility of large language models and their limitations. Instruction must emphasize data provenance within the context of each discipline. In economics coursework, students should trace how data is generated, aggregated, and distorted. Leadership courses should examine how human performance is quantified and measured. Logistics instruction should focus on how supply chain data is collected and where it commonly degrades. Across all subjects, the unifying objective remains consistent. Operators must understand where data originates, how it is collected, and why it can mislead even when it appears authoritative.

The second tier builds upon this foundation by integrating data literacy into the core of operational planning at intermediate PME levels. This phase parallels the approach adopted by the Army Logistics University, which treats data not as an adjunct to planning but as an integral component of every operational function. Officers must move beyond recognizing data sources and begin incorporating data-driven insights into complex decision-making under uncertainty. At this stage, instruction must confront the ethical and adversarial dimensions of AI directly. Planners should examine concrete cases of responsible data use, algorithmic bias, and the legal implications of data-informed decisions. More importantly, they must learn how adversaries deliberately weaponize data. By studying how intelligence feeds are manipulated or how the signatures of military systems are altered to deceive algorithms, mid-level officers develop the habit of viewing data through an adversarial lens. That habit is essential to building plans that remain resilient when AI encounters an opponent intent on exploiting its vulnerabilities.

The third and most advanced tier applies to senior PME and shifts the instructional focus from recognizing adversarial manipulation to managing its strategic risk. Senior leaders must remain fluent in the unethical and adversarial uses of data introduced at the intermediate level, but executive education must prioritize judg-

ment rather than technique. At this stage, commanders must learn to weigh the risks associated with AI integration across major and theater-level operations. The curriculum must prepare them to assess the strategic implications of corrupted or incomplete data, balance the operational necessity of autonomous systems against the risk of adversarial exploitation, and decide when to rely on AI-derived courses of action when consequences are irreversible. These decisions do not admit technical certainty. They demand informed risk acceptance grounded in an understanding of how AI fails as well as how it succeeds.

Responsibility for mandating and enforcing this progression rests with Joint Staff J-7, which serves as the proponent for both the Officer Professional Military Education Policy under Chairman of the Joint Chiefs of Staff Instruction 1800.01G²⁶ and the Enlisted Professional Military Education Policy under CJCSI 1805.01C.²⁷ This position grants J-7 authority over curriculum requirements spanning the full continuum of PME, from initial enlisted instruction through senior-level development. The Enlisted Professional Military Education Policy already directs leaders to think critically, anticipate and mitigate risk, and operate effectively in complex environments. These attributes align directly with the data literacy competencies advanced in this article. The remaining deficiency is structural. Unlike officer Joint PME, enlisted Joint education is not mandated by statute and therefore functions as guidance rather than requirement. Codifying data literacy as a defined educational outcome under both policy instruments would close this gap and establish the competency as a force-wide standard rather than an institutional aspiration.

A Risk Framework for AI Integration, The Acceptable Level of Risk Schema

The preceding sections identify a central problem. AI is moving rapidly into consequential military applications at a pace that exceeds the institution's ability to produce operators capable of validating its outputs. The DOW currently lacks a shared decision framework that allows commanders to assess, articulate, and communicate the degree of risk they assume when employing AI systems, particularly in conditions that strain or exceed the limits of human verifiability and technical competence.

²⁶ Chairman of the Joint Chiefs of Staff, *Officer Professional Military Education Policy*, CJCSI 1800.01G (Washington: Joint Staff, 15 April 2024), <https://www.jcs.mil/>.

²⁷ Chairman of the Joint Chiefs of Staff, *Enlisted Professional Military Education Policy*, CJCSI 1805.01C (Washington: Joint Staff, 1 November 2021), <https://www.jcs.mil/>.

Such a framework already exists within military practice. The Acceptable Level of Risk (ALR) construct, codified in Chairman of the Joint Chiefs of Staff guidance and embedded across service operational risk management doctrine, provides a five-tier structure through which commanders routinely evaluate risk to the force and to the mission: low, moderate, significant, high, and extreme. Commanders use this framework to ensure that risk acceptance occurs deliberately and at the appropriate level of authority. Because the model is familiar to anyone who has participated in operational planning or mission execution, it offers a natural foundation for addressing a new category of risk. That category is not the external threat posed by an adversary, but the internal risk inherent in decisions derived from AI.

Risk associated with AI integration turns on two interacting variables. The first is verifiability, defined as the degree to which a human operator can trace an AI-generated output to its source data, assess the logic that produced it, and independently confirm its validity before acting. This capability proves especially critical in detecting data poisoning, manipulation, or model degradation. The second variable is consequence, defined by what occurs if the AI output is wrong and a human proceeds on that basis. Together, these factors delineate what this article terms the verifiability gap. That gap measures the distance between what an AI system asserts and what an operator can actually confirm. At one end of the spectrum, the output remains fully verifiable and errors are correctable. At the other end, human inspection disappears altogether and machine-generated outputs drive irreversible action. Each tier of the ALR framework corresponds to a distinct position along this gradient.

Applying the Schema

The value of this framework lies not in classification for its own sake, but in the creation of a shared vocabulary for operational risk. The schema links the data integrity vulnerabilities to the platform integration challenges examined in earlier portions of this article and the educational shortfalls described above. It allows commanders to articulate the nature of the risk they accept when relying on AI-derived outputs, rather than defaulting to implicit trust or informal judgment. When a commander employs an AI-enabled targeting, intelligence, or weaponeering system without organic expertise to evaluate the underlying sensor fusion and training data, the decision falls into the High tier. In such circumstances, the probability of catastrophic failure increases substantially, carrying the potential for significant materiel loss and severe military or civilian casualties.

The Five-Tier ALR Schema

ALR Level	Verifiability and Procedures	Operational Implications and Examples
Low	The output is fully explainable. The operator can see the source data, trace the system's reasoning, and independently verify the conclusion before taking action. There is no ambiguity in how the answer was reached.	An AI tool generating maintenance items from aircraft health monitoring data, where a crew chief independently reviews each item before signing off, operates at this tier. The operator retains full visibility into the source data and can identify and correct any errors before action is taken.
Moderate	The output uses verifiable data, but full independent verification is not feasible due to time constraints or the volume of information. The operator works within their domain of expertise and can identify anomalies on inspection.	Domain knowledge serves as a natural filter. An intel analyst reviewing an AI-generated edge-processed target summary against known pattern-of-life data before passing a recommendation to a strike cell operates at this tier. A logistics officer evaluating an AI-generated sustainment plan for a mission type they have executed before is a second example.
Significant	The operator or team has only surface-level knowledge of the domain. They cannot reliably catch specific errors on inspection, but have enough familiarity to recognize obviously wrong outputs. Verification is partial.	A joint fires officer integrating an AI-generated targeting recommendation into a dynamic targeting cycle, without organic ISR expertise to audit the underlying sensor data, operates at this tier. The output may appear plausible, but specific errors in the source data are unlikely to be caught before action is taken.
High	The operator or team is outside their knowledge space entirely and lacks the expertise to evaluate the output, but they closely monitor the material to the best of their ability. Manipulation or distributional shifts could go undetected without a meaningful human filter for error.	A manned wingman accepting a CCA targeting solution in a degraded communications environment, where the onboard AI's sensor data cannot be interrogated at execution, operates at this tier. A strike planner acting on a targeting product from a sensor architecture whose training data provenance is unknown to anyone in the current chain of command is a second example.
Extreme	Full autonomy with zero verification. The system acts, or its output drives action directly, with no human check at any point in the chain.	Kinetic, cyber, or space effects are executed solely on machine judgment with no human intervention or validation. Technical capability to remove the human from the loop is an engineering minor update; the policy is the only limiting factor.

A similar dynamic applies when forward crews act on edge-processed threat assessments generated by models they cannot interrogate because the relevant training data remains classified and controlled by a contractor. In those cases, the appropriate risk tier falls between Significant and High, depending entirely on whether anyone in the decision chain possesses sufficient domain expertise to assess the output critically. When a CCA executes an autonomous engagement authority package in a degraded communications environment without the possibility of realtime human verification, the decision reaches the Extreme tier. This

classification applies regardless of formal policy language asserting human authority over weapons release. Authority that cannot be exercised in practice does not meaningfully constrain risk.

The framework does not instruct commanders to avoid accepting Significant or High levels of risk. Accepting such risk is inherent to command and occurs routinely in military operations. The framework instead requires that risk be identified, communicated, and consciously accepted at the appropriate level of authority. In this respect, it mirrors existing doctrine governing operational risk management, which does not eliminate danger but ensures accountability for its acceptance.

An institution that requires commanders to state explicitly which ALR tier governs their use of AI-derived products establishes an accountability mechanism that does not currently exist. It creates incentives to reduce risk over time by developing expertise that moves decisions from High to Significant, and from Significant to Moderate. It also provides a practical foundation for PME. Teaching operators to identify their position relative to AI systems, and to recognize the limits of their own ability to verify machine-generated outputs, captures the operational meaning of data literacy. Naming the risk becomes the precondition for managing it.

Conclusion: Naming the Gap

This article does not argue that AI is inherently dangerous. AI constitutes a powerful and indispensable capability and will shape the future of warfare regardless of institutional preference. The danger lies elsewhere. It lies in the widening gap between what AI systems can generate and what the institution can meaningfully verify. That verifiability gap is expanding at a pace that exceeds the military's ability to adapt its organizational structures, educational models, and risk frameworks to control it.

Evidence of this gap appears consistently across each domain examined. In data governance, the DOW cannot reliably identify its AI workforce or systematically trace the provenance of the training data upon which critical platforms depend. Legislative action has acknowledged the corresponding vulnerability, explicitly recognizing adversarial efforts to corrupt and poison that data. In operations, the YFQ-44A has already progressed to weapons integration testing, while autonomy software architectures deliberately separate algorithms from hardware to enable rapid updates. At the same time, analysis by the Center for Strategic and International Studies (CSIS) has shown that the engineering distance between semi-autonomous and fully autonomous weapons systems is far smaller than the

policy distance intended to regulate them.²⁸ In education, the DOW still lacks a coherent model for developing an AI-ready force at scale, even as it commits hundreds of millions of dollars to accelerating AI capability. PME has not yet incorporated data literacy as a core operational skill commensurate with the authority increasingly delegated to machines.

None of these deficiencies reflects a lack of ambition. The Department of the Air Force is moving with visible urgency. Recent generative AI initiatives, sustained engagement with firms such as Anthropic, and the department's expanding relationship with OpenAI all underscore a serious institutional commitment to accelerating capability. The CCA program continues to move aggressively through development, and the Joint All-Domain Strategist program seeks to produce officers capable of operating across domains while identifying data inconsistencies as an asymmetric advantage. Investment in AI capability is real, deliberate, and consequential. The problem is not momentum. It is an imbalance. Investment in technology has decisively outpaced investment in the human infrastructure required to employ that technology with discipline and restraint.

The argument presented in this article converges on a single conclusion. Data literacy is the connective tissue linking technical capability, operational trust, and institutional readiness. An operator who understands where data originates, how it degrades, and how to recognize the signatures of a model operating beyond its training parameters is not merely a more competent user of AI. That operator embodies the professional judgment required to distill clarity from complexity. More importantly, this capability enables commanders to make life-and-death decisions based on information produced by autonomous systems operating at the tactical edge. The skill itself does not vary by platform or domain. What varies is the institutional recognition of its importance. Until data literacy is treated as a foundational competency rather than a supplemental one, the DOW will continue to field advanced systems without building the capacity to govern their use at scale.

The proposed five-tier ALR framework offers one concrete means of closing the gap identified in this article. Its purpose is not to impose another administrative layer on an already overburdened force. It provides commanders with a common language for a category of risk that often goes unnamed in operational and planning environments. An operator who can state explicitly that a particular decision entails a High-tier level of AI risk because it lies outside their domain of verification has already completed the most important intellectual task. Naming

²⁸ Gregory C. Allen, "DoD Is Updating Its Decade-Old Autonomous Weapons Policy, but Confusion Remains Widespread," Center for Strategic and International Studies, 6 June 2022, <https://www.csis.org/>.

the risk is the necessary precondition for managing it. Unacknowledged risk cannot be controlled.

The broader policy implications follow directly from this logic. The DOW must define data literacy as a formal competency, supported by explicit training standards, billet coding, and PME requirements, in the same manner used to govern other mission-critical skills. The analytical groundwork already exists. What remains absent is an institutional commitment to treat data literacy not as an elective subject or an ancillary technical topic, but as a foundational requirement for those who operate, plan alongside, and ultimately command increasingly autonomous systems.

Throughout its history, the military has confronted the persistent tension between the speed of technological change and the capacity of institutions to absorb it responsibly. That tension does not resolve through technology alone. It resolves when leaders identify the problem, construct a framework to manage it, and demand institutional adaptation equal to the consequence at stake. The AI integration gap outlined in this article warrants that level of seriousness. The issue is not that AI poses an inherent threat. It is that an institution unable to verify what its own machines are doing has surrendered a critical measure of control over its decision-making. In war, that loss of situational awareness carries strategic and human consequences that no degree of technical sophistication can offset. 🦅

Lt Col Ryan Blakeney, USAF

Lieutenant Colonel Blakeney is a US Air Force RQ-4B weapons officer with extensive experience in high-altitude intelligence, surveillance, and reconnaissance operations. He currently serves as a Joint All Domain Strategist Instructor at Air Command and Staff College, Maxwell Air Force Base, Alabama. His research focuses on the integration of artificial intelligence into wargaming, with particular emphasis on data integrity, operational trust, and risk management frameworks for AI-enabled military systems.