NEWS UNITED STATES AIR FORCE

Educating the AI-Ready Warfighter: A Framework for Ethical Integration in Air Force Professional Military Education

Published June 17, 2025

By Lt. Col. Christopher J. Smith, USAF

Finalist in the Spring AI and Military Education Innovation Essay Contest

Artificial intelligence (AI), particularly large language models (LLMs), is rapidly reshaping the U.S. Air Force's educational and operational environment. As these tools become more capable and accessible, their integration into Professional Military Education (PME) presents both a strategic opportunity and an ethical challenge. LLMs offer substantial benefits in streamlining instruction, enhancing critical thinking, and simulating complex scenarios, but they also raise concerns related to data security, bias, and the erosion of human judgment.

Effective integration of AI into PME demands more than adoption; it requires clear governance, transparency, and faculty preparedness. Without thoughtful policies, AI use could compromise creativity, diminish intellectual diversity, or unintentionally reinforce systemic biases. At the same time, if implemented responsibly, AI can augment leadership education and improve warfighting readiness by equipping officers with tools for analysis, planning, and decision support.

This paper examines the boundaries between AI assistance and human judgment, identifies potential biases and limitations in AI-generated educational content, and explores the risks of overreliance on AI within PME. It proposes a tiered governance framework and outlines pathways for securely scaling AI integration through policy innovation, faculty development, and operational test environments. By approaching AI as a cognitive amplifier—rather than a substitute for leadership—PME can responsibly harness its potential to prepare future military leaders for the evolving battlespace.

Ethical and Governance Implications of Integrating AI in Air Force PME

Integrating AI tools like large language models into Air Force PME is more than a matter of technological adoption; it is a transformative shift that compels a re-evaluation of existing ethical norms and governance frameworks, particularly in handling sensitive or classified information. While LLMs promise substantial benefits in personalized learning, administrative efficiency, and critical thinking support, their application within a military context requires stringent ethical oversight and tailored governance mechanisms to address specific confidentiality, security, and institutional trust risks.

One of the central governance challenges in integrating LLMs into Air Force PME is managing the intersection between these AI systems and classified or sensitive information. Most commercially available LLMs operate on open-access infrastructure and are not certified for handling classified material.[1] If educators or students inadvertently input sensitive or classified information into an LLM like ChatGPT, it could result in a serious security breach. Therefore, PME institutions must develop clear, enforceable governance policies that delineate the boundaries of LLM usage in classified environments. They must also be aware of classified data that should not be entered into Generative AI or LLMs unless the LM has been cleared and approved by the institution's security department.[2]

The ethical integration of AI into PME hinges on well-articulated acceptable use policies. Overly broad or restrictive guidelines—such as those stating, "AI use is forbidden in the production of this paper"—risk being unenforceable and counterproductive.[3] Instead, policies should define key terms such as "generative AI," "training data," and "LLM" and provide clear examples of acceptable and impermissible uses. Additionally, transparency must be a core principle: Students and faculty should be required to disclose AI use in academic and research contexts.[4]

AI's integration into PME presents the ethical challenge of delineating the boundary between machine assistance and human judgment. While LLMs can aid in generating hypothetical ethical dilemmas or summarizing historical military conflicts, they cannot make morally sound decisions in high-stakes contexts like combat operations or rules of engagement. Furthermore, the inherent unpredictability and unreliability of current LLMs pose additional challenges. Their tendency to produce unexpected or erroneous outputs makes them unsuitable for tasks requiring high accuracy and trust, such as military decision-making and education. The lack of transparency in how LLMs generate responses further complicates their integration into critical fields like PME.[5]

A critical governance issue is addressing the current skill gap among educators regarding AI use. Many PME instructors may resist LLM integration due to unfamiliarity or fear of undermining traditional educational methods. A lack of understanding about how LLMs work often leads to hesitation or outright rejection of their use in the academic setting.[6] To close this gap, PME institutions must implement robust faculty development programs that include training on AI functionality, ethical implications, and operational boundaries, particularly concerning the handling of sensitive information.

Long-term, the ethical integration of AI in PME must move beyond isolated course policies toward a cultural shift that embeds AI literacy and integrity into the fabric of military education. A culture of responsible AI usage is essential for reducing stigma and ensuring that AI tools are treated neither as forbidden magic nor infallible oracles.[7] Ultimately, ethical AI integration in PME demands a cultural shift beyond individual course policies to institutional norms of responsible, transparent use.

Where Do We Draw the Line Between AI Assistance and the Need for Human Judgment in Developing Future Military Leaders?

The line between AI assistance and human judgment in developing future military leaders must be drawn at the boundary between *support* and *substitution*. Artificial intelligence, particularly large language models, can play a valuable role in augmenting education, offering learners ways to explore counterarguments, simulate scenarios, and analyze doctrinal concepts. These tools act as cognitive amplifiers, enhancing comprehension and fostering intellectual engagement. However, AI must not replace the reflective, ethical, and strategic reasoning at the core of military leadership.

AI can assist in generating hypothetical dilemmas or summarizing complex texts, but it lacks the moral reasoning, emotional intelligence, and experiential understanding required for leadership decisions. The model's suggestions are rooted in patterns—not original thought. AI does not 'understand' in the human sense; it predicts plausible text based on prior data. It lacks intuition, empathy, and moral sensibility.[8]

Military leadership involves analyzing data and applying judgment under uncertainty—balancing risks, navigating ethical dilemmas, and bearing the weight of accountability. These require human traits such as moral courage, emotional intelligence, empathy, and experience—traits that no algorithm can replicate.[9]

Moreover, military leadership involves accountability. No matter how refined an AI system becomes, individuals are responsible for decisions, especially in combat or high-stakes diplomatic situations. This moral burden cannot be outsourced to machines.[10]

Thus, AI's role in PME should be limited to enhancing thought and not making decisions. Institutions must enforce this distinction through clear policies, ethical training, and revised assessment methods prioritizing how and why decisions are made—not just the conclusions reached. In doing so, PME can harness AI's strengths without compromising the human qualities essential to command.

What Biases Could AI Inadvertently Introduce into Educational Content and Assessments, and How Can We Mitigate or Correct Them?

Large language models are poised to revolutionize learning by streamlining instruction, enhancing accessibility, and supporting individualized education. However, the inherent limitations and training processes behind these systems pose risks that may inadvertently introduce bias into educational content and assessments, potentially undermining equity, accuracy, and intellectual development. Understanding the nature of these biases and implementing safeguards is essential for responsibly leveraging AI within military and higher education institutions.

AI systems, particularly large language models (LLMs), can inadvertently introduce several types of bias into educational content and assessments due to the nature of their training data and how they generate responses.[11] These biases include cultural, linguistic, and racial biases and overemphasizing dominant historical or political narratives. Because LLMs are trained

on vast amounts of internet-based and published content—much of which reflects Western norms and perspectives—they may underrepresent or mischaracterize minority viewpoints, non-Western ideologies, or non-standard English dialects. This can marginalize students from diverse backgrounds and reinforce stereotypes in military education environments. For example, studies have demonstrated that LLMs like ChatGPT show favoritism toward speakers of Standard American English (SAE), portraying them as more competent and trustworthy than those who speak dialects such as Kenyan English or Indian English.[12]

Additionally, LLMs may confidently produce false or misleading information, referred to as "hallucinations," due to limitations in predicting text based on patterns rather than factual verification. These inaccuracies can impact assessments and misinform students who lack the expertise to verify AI-generated claims.[13]

LLMs trained on mainstream sources may perpetuate dominant political or historical narratives while minimizing or omitting alternative viewpoints. In PME contexts, this could lead to an incomplete or biased understanding of past military conflicts, doctrines, or ethical debates.

As LLMs are a product of their "learning," there will always be a concern for bias and false information. To mitigate these issues, PME institutions must take a multi-pronged approach. First, they should promote AI literacy among students and faculty, helping users understand where bias can originate and how to critically evaluate AI-generated content.[14] Second, AI should support, not replace, human judgment in assessments. Faculty oversight is essential to detect inaccuracies, recognize diverse communication styles, and ensure fairness. Experienced faculty can often identify when AI misclassifies content based on dialect or writing style.[15] Third, developers of military-specific LLMs should prioritize more diverse and representative training datasets that include a range of cultures, dialects, and historical perspectives.[16] Finally, institutions must adopt clear ethical guidelines and regularly audit AI tools used in education to identify and correct biased behavior.[17]

AI tools hold tremendous promise in enhancing the reach and personalization of Professional Military Education. However, they are not neutral instruments. Bias—whether in the form of skewed perspectives, misinterpretation of dialects, or misinformation—can erode trust, diminish educational quality, and reinforce systemic inequities if not properly addressed. The path forward must involve deliberate strategies to mitigate bias, strengthen human oversight, and empower both faculty and students with the skills to critically engage AI. Only then can AI truly serve as an educational ally rather than a hidden adversary.

Does Reliance on AI in PME Risk Homogenizing Thought or Stifling Creative Leadership?

As artificial intelligence tools like LLMs become more prevalent in PME, questions about their long-term impact on student development have emerged. Among the most pressing is whether increased reliance on AI risks homogenizing thought and stifling the kind of creative, critical leadership necessary in complex military environments. While LLMs offer immense benefits in streamlining learning and expanding access to information, overdependence may indeed threaten the intellectual diversity, innovation, and independent reasoning that PME institutions are meant to cultivate.

In military education, where the goal is to prepare leaders for unpredictable, high-stakes environments, originality, critical thinking, and moral imagination are essential. If students begin to lean too heavily on AI for generating ideas, structuring arguments, or exploring scenarios, they risk becoming passive consumers of preprocessed information. This reliance can dull intellectual curiosity and discourage the kind of lateral thinking that leads to innovation on the battlefield or in strategy development.[18]

The risk is particularly pronounced when AI is used to brainstorm or draft arguments in assignments designed to stimulate original thought. Even if students edit or expand on AI-generated content, the initial framing—suggested by the model—can shape the direction of inquiry, limiting intellectual exploration. AI tools might subtly nudge users toward conclusions that are already "common sense" or highly represented in their training data, thus curbing the pursuit of outlier or disruptive ideas that are often the source of innovation in military strategy and leadership.[19]

Moreover, AI models are not trained to challenge institutional assumptions or propose unconventional alternatives. They mirror the data they've been fed, much of which is rooted in past doctrine and mainstream perspectives. This can limit exposure to divergent viewpoints or emerging concepts that lie outside the model's dominant training scope.

Despite these risks, AI need not threaten creativity if PME institutions implement guardrails and pedagogical strategies that prioritize critical engagement. One approach is to integrate AI literacy into the curriculum, teaching students not just how to use

AI but also how to question and critique its outputs. Exercises where students compare AI-generated responses to their own or identify biases in AI-generated scenarios can foster deeper reflection and analytical rigor.[20]

As discussed, transparency is another key. To remove the stigma of AI use, students should be encouraged to document how AI tools were used in an assignment to help educators discern where AI may be narrowing or shaping thought. This also reinforces academic honesty and helps students take ownership of their intellectual contributions, skills that will be required by warfighters to explain the reasoning behind their decisions.

Ultimately, while AI can support the learning process, PME must ensure that it does not replace the rigorous, creative, and often uncomfortable thinking required of future military leaders. Thoughtful integration, not blind adoption, is key to preserving the cognitive diversity and leadership adaptability that national defense demands.

Strengthening the Operational Link Between PME and the Warfighting Force Through AI Integration

AI integration must bridge academic development and operational application to ensure that Professional Military Education contributes directly to warfighting readiness. Officers trained in the ethical and practical use of large language models gain more than theoretical knowledge—they acquire tools that enhance situational awareness, critical thinking, and planning efficiency in operational environments. By engaging with AI during PME, officers become familiar with using LLMs to simulate doctrinal scenarios, assess variables, or test strategic hypotheses. These academic exercises mirror real-world applications, allowing officers to rehearse decision-making under complexity before encountering it in the field.

PME also plays a critical role in shaping leaders who are not only AI-literate but strategically adept at translating AI capabilities into combat power. AI-ready leaders are those who understand how to responsibly incorporate machine assistance into planning cycles, wargames, and intelligence processes while preserving human judgment as the ultimate authority. This is particularly important given AI's limitations. While LLMs can provide summaries, generate plausible courses of action, or flag logistical inconsistencies, they cannot intuit geopolitical nuance or exercise moral judgment. PME curricula that emphasize critical evaluation of AI outputs, bias detection, and ethical constraints are essential for developing leaders who can clearly and confidently command an AI-augmented battlespace.

LLMs also offer direct value to the warfighter. In mission planning, AI can help synthesize doctrine, environmental data, and adversary patterns into coherent operational plans. In logistics, predictive AI models can assist with sustainment forecasting, supply chain optimization, and rapid re-tasking based on battlefield dynamics. For intelligence, surveillance, and reconnaissance (ISR) analysis, LLMs can rapidly parse text-based intel, summarize patterns, and surface emerging threats from large data sets. These functions do not replace human analysts but extend their reach and accelerate insight generation.

Toward a Tiered Policy Framework and Scalable AI Implementation

While this paper highlights the importance of ethical oversight and governance in integrating AI into Professional Military Education, a more robust policy roadmap is essential to ensure secure, scalable, and mission-aligned adoption across the Air Force. A tiered AI policy framework should be implemented to govern usage across varying levels of information sensitivity. Commercially available LLMs like ChatGPT may be used in unclassified environments for ideation, editing, and general educational support—provided users disclose AI involvement. CUI material, only government-approved AI tools hosted within secure cloud environments with identity tracking and audit logs should be permitted. In classified settings, AI use should be restricted to models specifically trained, hosted, and accredited within Department of Defense (DoD) networks, ensuring zero risk of data exfiltration.

To operationalize these policies, PME institutions should adopt sandbox environments that allow students and faculty to experiment with AI tools under controlled conditions. These "testbeds" would simulate real-world applications—such as AI-supported wargames, ISR summarization, or logistics planning—without risking exposure of sensitive data. Such environments can also serve as incubators for developing military-specific AI tools while collecting feedback to inform broader implementation efforts.

Innovation units such as AFWERX, SAF/AI, and the Air Force Research Laboratory (AFRL) should be pivotal in scaling AI integration. AFWERX can accelerate prototype development of secure LLMs tailored to PME needs. SAF/AI should lead in policy development, compliance oversight, and the accreditation of AI tools for military use. Meanwhile, AFRL could assist in validating

the operational utility of AI applications identified through PME test programs. Collaboration across these units would ensure that AI adoption is not only technically feasible but also aligned with the strategic and ethical imperatives of modern warfare.

Conclusion

Ultimately, PME must serve as a proving ground for these applications by equipping officers to understand AI and to wield it in support of national defense. The integration of AI tools like LLMs into Air Force PME is not merely a technological evolution—it is a strategic imperative that demands thoughtful policy, ethical vigilance, and operational foresight. While AI can enhance critical thinking, streamline instruction, and support warfighting readiness, its misuse or overreliance could compromise judgment, creativity, and data integrity. Therefore, the Air Force must adopt a proactive and phased approach to AI governance in education, training, and operations.

To guide implementation, the following roadmap outlines key steps the Air Force should take:

Now:

- Publish interim guidance on acceptable AI use within PME, emphasizing transparency, ethical constraints, and classification restrictions. Avoid generalized restrictions of AI on assignments.
- Launch faculty and student AI literacy training programs across PME institutions.
- Designate pilot PME institutions to serve as innovation testbeds for AI-enhanced learning.

Short-Term (6–18 months):

- Develop and adopt a tiered AI policy framework distinguishing rules for unclassified, CUI, and classified environments.
- Establish secure sandbox environments for PME to test AI tools under realistic but non-operational conditions.
- Engage AFWERX, SAF/AI, and AFRL in supporting PME-based experimentation and feedback loops for operational AI use cases.

Long-Term (18 months and beyond):

- Scale validated AI tools and use policies across PME and operational training pipelines.
- Integrate AI into joint exercises and wargaming, including adversarial simulation and logistics forecasting.
- Create a PME AI Center of Excellence to lead curriculum development, policy evolution, and ethical AI research.

This roadmap ensures that AI becomes not just a pedagogical enhancement but a foundational capability in preparing warfighters for an increasingly complex battlespace—where human judgment remains irreplaceable, but AI stands as a powerful force multiplier.

Lieutenant Colonel Christopher Smith is assigned to the Air War College, Air University, Maxwell AFB, AL. Lt Col Smith received a direct commission in the U.S. Air Force Dental Corps following graduation from The University of Maryland, Baltimore College of Dental Surgery. He completed his residency in oral and maxillofacial surgery, is a Diplomate of the American Board of Oral and Maxillofacial Surgery, and a diplomate of the National Dental Board of Anesthesiology. Lt Col Smith currently serves as the U.S. Air Force delegate for the American Dental Association and alternate delegate for the American Association of Oral and Maxillofacial Surgeons and is a Fellow of the International College of Dentists. Lt Col Smith also holds a Master of Education in Health Professions Education from Uniformed Services University and works with multiple residency programs around the country on their curriculum and educational processes.

---

[1] Jon Martindale, "Unlocking Llama 3's Potential: What You Need to Know," *Lifewire*, October 14, 2024.

[2] "Artificial Intelligence Policies: Guidelines and Considerations," *Duke Learning Innovation & Lifetime Education* (blog), accessed March 12, 2025; Garcia Odom, "Department of Homeland Security: Commercial Generative Artificial Intelligence Interim Rules of Behavior" (Department of Homeland Security, November 28, 2023).

[3] "Artificial Intelligence Policies"; Stephen Renner, David Palkki, and Nicholas Sambaluk, "Foundations of Strategy AY25 Syllabus" (Air War College, July 24, 2024); David J. Gordon and John P. Geis II, "Air, Space, and Cyber Power in Future Conflict Course (AY 2025)" (Air War College, July 26, 2024).

[4] Miguel A Cardona, Roberto J Rodríguez, and Kristina Ishmael, "Artificial Intelligence and the Future of Teaching and Learning: Insights and Recommendations" (United States: Office of Educational Technology, May 2023), 49.

[5] James Fodor, "Intrinsic Limitations of GPT-4 and Other Large Language Models, and Why I'm Not (Very) Worried about GPT-n," *Effective Altruism Form*, June 3, 2023.

[6] Nikolas McGehee, "Breaking Barriers: A Meta-Analysis of Educator Acceptance of AI Technology in Education | Michigan Virtual," *Michigan Virtual*, November 21, 2024.

[7] Simon Larsson and Martin Viktorelius, "Reducing the Contingency of the World: Magic, Oracles, and Machine-Learning Technology," *AI & SOCIETY* 39, no. 1 (February 2024): 2.

[8] Yiheng Liu et al., "Understanding LLMs: A Comprehensive Overview from Training to Inference" (arXiv, January 6, 2024).

[9] Kate Robinson, Bernard McKenna, and David Rooney, "The Relationship of Risk to Rules, Values, Virtues, and Moral Complexity: What We Can Learn from the Moral Struggles of Military Leaders," *Journal of Business Ethics* 179, no. 3 (September 1, 2022): 749–66.

[10] James Kraska, "Command Accountability for AI Weapon Systems in the Law of Armed Conflict," *International Law Studies* 97, no. 407 (2021): 409.

[11] Jinsook Lee et al., "The Life Cycle of Large Language Models in Education: A Framework for Understanding Sources of Bias," *British Journal of Educational Technology* 55, no. 5 (2024): 1982–2002.

[12] Ritwik Gupta et al., "Linguistic Bias in ChatGPT: Language Models Reinforce Dialect Discrimination," *The Berkeley Artificial Intelligence Research Blog*, September 20, 2024.

[13] David E. Balch and Robert Blanck, "Mitigating Hallucinations in LLMs for Community College Classrooms: Strategies to Ensure Reliable and Trustworthy AI-Powered Learning Tools," *Faculty Focus | Higher Ed Teaching & Learning*, May 15, 2024.

[14] Aimé Anita and Kiu Publication Extension, "Digital Literacy in the 21st Century: Preparing Students for the Future Workforce," *Research Output Journal of Arts and Management* 3, no. 3 (August 30, 2024): 40–43.

[15] Ahmed M. Elkhatat, Khaled Elsaid, and Saeed Almeer, "Evaluating the Efficacy of AI Content Detection Tools in Differentiating between Human and AI-Generated Text," *International Journal for Educational Integrity* 19, no. 1 (December 2023): 1–16.

[16] Dmitry I. Mikhailov, "Optimizing National Security Strategies through LLM-Driven Artificial Intelligence Integration" (arXiv, May 7, 2023).

[17] Daniel Birks and Joseph Clare, "Linking Artificial Intelligence Facilitated Academic Misconduct to Existing Prevention Frameworks," *International Journal for Educational Integrity* 19, no. 1 (December 2023): 1–10.

[18] Chunpeng Zhai, Santoso Wibowo, and Lily D. Li, "The Effects of Over-Reliance on AI Dialogue Systems on Students' Cognitive Abilities: A Systematic Review," *Smart Learning Environments* 11, no. 1 (June 18, 2024): 28.

[19] Tom Duenas and Diana Ruiz, "The Risks of Human Overreliance on Large Language Models for Critical Thinking," November 12, 2024.

[20] Anita and Extension, "Digital Literacy in the 21st Century: Preparing Students for the Future Workforce."