

Understanding Department of War Responsible AI Needs & Requirements

An Examination of DoW's AI Ethics and Policies for Trustworthy
Adoption of AI-Enabled Systems



Logan O. Mailloux, PhD, CSEP, CISSP (retired US Air Force Lt Col), logan.o.mailloux@lmco.com
Vincent M. Santore, MS, MBA, vincent.m.santore@lmco.com

Our Goals And Motivations For This Presentation:

1. Answer the question: “why does this topic deserve attention now?”
 - Ukrainian drone’s autonomous kill decision
 - Various State of the Art (SOTA) models have rapidly expanded their capability and have broken through intended safeguards
 - Anthropic’s Mythos Model & OpenAI’s Hugging Face incident
2. Increase our understanding of the DoW’s three primary guiding documents in their current forms
 - Discuss the guiding documents and standards, as well as gaps in their current and future needs
3. Discuss the “future requirements” our industry must be prepared for which have evolved faster than our guiding documents
4. Provide suggestions to increase trust between AI powered systems and military operations



2026 has been a year of incredible growth, as well as uncertainty

Where We Are Currently: 3 Guiding Documents

The Defense Industry currently has 3 high-level guiding documents regarding “Trustworthy AI” usage

1. Department of War Directive 3000.09 “Autonomy of Weapon Systems” (2012, Updated 2023, Currently in Revision)
 2. Department of War’s (DoW’s) Five Principles of AI Ethics (2020)
 3. Office of Management and Budget Presidential Memo 25-21 “Accelerating Federal Use of AI through Innovation, Governance, and Public Trust” (2025)
- Bonus Document: National Security Presidential Memorandum – 11 (2026)
 - Additionally, the Federal government continues to push for AI to be regulated at the Federal, rather than state level, but no single comprehensive legislation or guiding document currently exists



Note: the Law of Armed Conflict (LOAC) remains the basis for all DoW documents, guidance, and directives

DoW Directive 3000.09 “Autonomy of Weapon Systems” (2012, Updated 2023, Currently in Revision)

- Purpose of 2023 Revision:
 - Establishes policy and assigns responsibilities for developing and using autonomous and semi-autonomous functions in weapon systems, including armed platforms that are remotely operated or operated by onboard personnel.
 - Establishes guidelines designed to minimize the probability and consequences of failures in (...) weapon systems that could lead to unintended engagements.
 - Establishes the Autonomous Weapon Systems Working Group.
 - Currently under revision with updates and clarifications ongoing
1. Humans must remain in control at all times
 - Even when a system can act autonomously, a human must retain meaningful oversight (Humans in/on the loop) before lethal force is applied
 2. Legal compliance
 - All automated weapon systems must be demonstrably consistent with Laws of Armed Conflict, Internal Humanitarian, and U.S. statutory law
 3. Rigorous testing and validation are required
 - Continuous monitoring at every stage to ensure reliability, predictability, and ethical behavior
 4. Clearly defined accountability structures must be established at every stage of the AI usage, from development, to command, to use by the operators
 5. Transparency and documentation enable oversight, post-event analysis, and public trust

DoW's Five Principles of AI Ethics (2020)

1. Responsible

- DoW personnel will exercise appropriate levels of judgment and care while remaining responsible for the development, deployment and use of AI capabilities.

2. Equitable

- The department will take deliberate steps to minimize unintended bias in AI capabilities.

3. Traceable

- The department's AI capabilities will be developed and deployed such that relevant personnel possess an appropriate understanding of the technology, development processes and operational methods applicable to AI capabilities, including with transparent and auditable methodologies, data sources and design procedures and documentation.

4. Reliable

- The department's AI capabilities will have explicit, well-defined uses, and the safety, security and effectiveness of such capabilities will be subject to testing and assurance within those defined uses across their entire life cycles.

5. Governable

- The department will design and engineer AI capabilities to fulfill their intended functions while possessing the ability to detect and avoid unintended consequences, and the ability to disengage or deactivate deployed systems that demonstrate unintended behavior.

The DoW additionally released the *Responsible Artificial Intelligence (RAI) Toolkit*

OMB Presidential Memo 25-21 “Accelerating Federal Use of AI through Innovation, Governance, and Public Trust” (2025)

- High-Level Goals:
 - “Agencies must remove barriers to innovation and provide the best value for the taxpayer”
 - “Agencies must empower AI leaders to accelerate responsible AI adoption”
 - “Agencies must ensure their use of AI works for the American people”
 - Fostering public trust in the federal use of AI
- Required specific changes in workflow and process, moving from theoretical policy into tactile implementation, with given tasks for each Executive Branch Department and Agency

Organizational & People	Technological & Process
Named Chief AI Officer (CAIO) for each agency, who leads each Agency AI Governance Board	Develop a documented AI Strategy for identifying and removing barriers from mature, enterprise-wide AI use
Leverage training and resources to upskill existing staff, promote AI talent, and ensure accountability	When possible, leverage reuse in tools, data, and procedures among departments and agencies
Recruit, hire, train, retain, and empower an AI-ready workforce	Provide guidance on High-Impact AI capabilities and required testing
Achieve AI literacy for non-practitioners involved in AI	Mandate for continuous compliance plans every two years

Bonus: National Security Presidential Memorandum – 11 (2026)

- Artificial intelligence (AI) will be among the most transformative technologies to national security in the history of the United States. When adopted appropriately, AI can help protect our warfighters during peacetime and on the battlefield, enable precise operations that minimize harm to civilians, and ensure the United States continues to maintain technical overmatch against our adversaries and strategic competitors.
- The United States can and will responsibly accelerate the use of AI across intelligence and warfighting domains in line with American values. The United States possesses the most effective and moral military in the history of world. It is also among the most trusted institutions in American life. That trust is rooted in an unbroken chain of command and accountability, from our democratic process through civilian and military leadership, to the men and women who carry out the mission.

“The United States can & will responsibly accelerate the use of AI across intelligence & warfighting domains”

Recommendations

1. Re-examination of the five ethical AI principles for *Today's AI Capabilities*
2. Increase defense industry participation on AI related policies
 - “Don’t write policy about us, without us”
3. Keep AI compliance plans up-to-date with technology
 - Move to event based compliance (changes in operational context, changes or upgrades of model), rather than classic calendar-based approach
4. Building trust between military operators and high-consequence AI capabilities
 - Further research into understanding HITL / HOTL architectures
 - Recognize that HITL does not solve all our trust problems
5. Encourage national-level disclosure and agreement on the Use of AI
 - AI Models are improving exponentially, not linearly – “AI is an amplifier” for good or evil

“The nation that successfully implements AI principles will lead in AI for many years”

Questions?



References & Additional Sources

1. <https://spectrum.ieee.org/autonomous-drone-warfare>
2. <https://www.yahoo.com/news/world/articles/ukraine-ai-drones-entered-terminator-093000061.html>
3. https://www.army.mil/article/286795/lethal_unmanned_aerial_systems_safe_reliable_lethal_armaments_capability_at_every_echelon
4. <https://dora.dev/research/2025/dora-report/>

LOCKHEED MARTIN 