G1VE TECHNOLOGIES

# No Vendor Holds the Trigger: Why the Pentagon Is Rewriting Military AI Contracts

W.B. Carrington, MBA, LSSMBB, CTPRA, GRCA, GRCP

Mar 29, 2026

**Original source**

https://www.linkedin.com/pulse/vendor-holds-trigger-why-pentagon-rewriting-military-w-b--vdhue/

## Section I: The Fight Is Not Really About Anthropic. It's About Sovereignty in the Kill Chain.

The Pentagon's clash with Anthropic was never just a contract dispute between one AI lab and one government customer. It was a preview of a much bigger strategic confrontation:

- Who ultimately governs the use of artificial intelligence inside time-sensitive national security operations?

The Department of War has made its answer unmistakably clear. It does not intend to let any single vendor, however capable, insert private terms of service into the operational bloodstream of the Joint Force in a way that could constrain lawful Mission Planning, ISR-Enabled Targeting Support, Battle Management, or Mission Assurance in real time.

That matters far beyond one company. The Joint Force is moving toward a battlespace defined by compressed timelines, distributed operations, JADC2-style data fusion, and Mission Command under conditions of contested communications and decision overload. In that environment, AI is not just a support tool for PowerPoint briefings and back-office analytics. It is becoming part of the operating architecture through which commanders interpret the battlespace, assign resources, sequence actions, and preserve operational tempo. If that architecture is vulnerable to abrupt vendor-imposed constraint, then the United States has created a hidden dependency inside the very systems meant to protect national decision advantage.

This is why the stakes are doctrinal, not merely contractual. Mission command presumes decentralized execution aligned to commander's intent. That is hard enough when humans and legacy systems are involved. Add commercial AI models, proprietary safety stacks, opaque usage restrictions, and classified-network deployment realities, and suddenly the trigger is no longer only a weapons issue. It becomes a governance issue. Who can shape the last layer of decision support? Who can block it? Who monitors it? Who carries accountability when the system either overreaches or refuses to function? Those are no longer abstract policy questions. They are operational design questions now.

## Section II: Emil Michael 's "Holy Cow" Moment Exposed the Real Problem

The cleanest window into the Pentagon's thinking came from Office of the Under Secretary of War for Research and Engineering . In March 2026, he said he had a "holy, holy cow" moment when reviewing AI contracts already embedded in sensitive military commands. His concern was not that the Pentagon lacked AI access. His concern was that the access came wrapped in restrictions so sweeping that operators could be blocked from planning operations if those operations might potentially lead to kinetics. He also warned that if those contractual red lines were crossed, the model could theoretically "just stop in the middle of an operation." He specifically referenced commands responsible for air operations over Iran, China, and South America. That is not a software inconvenience. That is a mission assurance problem of the highest order.

That quote should be read slowly by anyone building AI for defense. The problem the Pentagon is signaling is not merely bias, hallucination, or cyber vulnerability. The problem is operational fragility created by contractual governance misalignment. If AI is used inside planning, intelligence support, command support, or high-speed decision environments, then a private vendor's restrictive terms can create a single point of mission failure even if the underlying model is technically sound.

Michael's further statement — "What we're not going to do is let any one company dictate a new set of policies above and beyond what Congress has passed" — is probably the most important sentence in the entire dispute. It reveals the Pentagon's emerging acquisition doctrine for AI: commercial firms may supply models, infrastructure, safety engineering, and domain-specific capabilities, but they do not get to supplant the legal and constitutional chain of authority over lawful military use. In short, the Department wants the benefits of frontier AI without ceding command prerogatives to unelected corporate actors.

For the defense market, that statement is a demand signal. The Pentagon is moving toward AI contracts that preserve sovereign operational control, support human accountability, and reduce the risk that mission-critical support will be interrupted by vendor policy conflicts at the moment of execution. That is not anti-safety. It is anti-fragility.

## Section III: The Anthropic Dispute Drew the Line Around Two Red Zones

To be fair, Anthropic's position was not random. Breaking Defense reported that the company refused to remove safeguards around two areas it believed could undermine democratic values:

- Mass domestic surveillance and fully autonomous weapons.

Anthropic CEO Dario Amodei said the company could not in good conscience accede to Pentagon demands to allow the model to be used for "all lawful purposes." The Pentagon, meanwhile, insisted it had no interest in using AI for either mass surveillance or autonomous weapons in violation of law, but objected to the idea that a private firm should define those limits in a way that exceeded statutory and policy boundaries.

This is exactly where the policy fight gets serious. Anthropic was arguing from an AI governance lens:

- Some capabilities should not be loosened because future misuse is too consequential.

The Pentagon was arguing from a constitutional and mission-command lens:

- Lawful military authority cannot be subordinated to a vendor's unilateral interpretation of acceptable use.

Both positions can sound reasonable in isolation. The conflict appears when AI moves from general-purpose assistance into operational relevance.

Then came the coercive pressure. Sean Parnell publicly demanded that Anthropic allow Pentagon use "for all lawful purposes," with a deadline and the warning that failure would trigger partnership termination and a "supply chain risk" designation. Anthropic said the Pentagon had also threatened to invoke the Defense Production Act to force the safeguards' removal. Whether or not every pressure tactic survives legal challenge, the strategic message was unmistakable:

- For mission-relevant AI, the Department of War is no longer willing to tolerate contractual ambiguity that could later become operational veto power.

Yet the story became even more revealing when Reuters later reported that a Pentagon memo allowed rare and extraordinary exemptions to Anthropic's phase-out when the capability was deemed mission critical and no viable alternative existed. The memo specifically indicated priority removal from systems supporting nuclear weapons and ballistic missile defense, while also acknowledging there could be waiver requests where operational reliance remained real. That carve-out says two things at once:

- First, the Pentagon wanted to reassert control; and,
- Second, it also recognized that AI dependence had already become deep enough that a clean severing was operationally difficult.

---

## Section IV: OpenAI's Deal Shows the Pentagon's Preferred End State

OpenAI's contrasting posture offers a useful view of where defense AI contracting may be headed. Reuters reported that OpenAI's classified-network deal with the Pentagon kept three explicit red lines:

- No mass domestic surveillance;
- No use to direct autonomous weapons systems; and,
- No high-stakes automated decisions.

It also retained a multi-layered safety model in which OpenAI preserved discretion over its safety stack, deployed through cloud-based infrastructure, and kept cleared OpenAI personnel in the loop. Shortly thereafter, Sam Altman said the agreement was also being modified to make clear that the services would not be used by intelligence agencies such as the NSA without a separate contract change.
That model is strategically significant because it looks like a negotiated middle path:

- Vendor-defined safety architecture is retained, but, operational usability is not gutted.
- Human beings remain in the loop.
- Certain red zones remain intact.
- The Pentagon gets deployable capability on classified networks.
- The vendor gets enforceable principles.

This is not a trivial compromise. It may well become the template for elite military AI contracting over the next several years.

For AI governance professionals, the deeper lesson is even more important. Effective defense AI governance will not come from pretending autonomy does not matter or from pretending risk can be solved by broad moral language. It will come from scalable governance mechanisms matched to degrees of autonomy, explicit support for human control and accountability, continuous monitoring and post-deployment oversight, defense-in-depth and containment strategies, and system-level risk assessments that examine the full operational environment rather than isolated model behavior. The OpenAI deal suggests the Pentagon is willing to live inside that structure. The Anthropic dispute suggests it is no longer willing to live outside it.

## Section V: What This Means for JADC2, Mission Assurance, and the Defense Industrial Base

This entire dispute should be understood through the lens of JADC2, ISR fusion, mission assurance, and resilient command. If the U.S. military is building a future in which sensor data from every domain is fused into common operational pictures and decision support is increasingly assisted by AI, then brittle AI contracting becomes a warfighting weakness. The problem is not just one model. The problem is allowing any indispensable node inside the digital decision layer to be governed in a way that is misaligned with operational continuity.

For the defense industrial base, this means the bar is rising. The Pentagon will increasingly prefer vendors that can provide:

- Mission-Usable AI under lawful Military Authorities;
- Robust Red-Team and Post-Deployment Oversight;
- Model Containment and Access Segmentation;
- Cleared Personnel and Secure Deployment Patterns;
- Auditable Human-Control Mechanism; and,
- Contractual Language that does not place National Security Missions at the mercy of abrupt vendor reinterpretation.

This also opens a space for a new class of integrators and governance specialists. The future winners may not be only frontier model labs. They may be the firms that can wrap those models in elite mission-assurance frameworks:

- Segmentation;
- Secure Orchestration;
- Policy Enforcement;
- Auditability;
- Fallback Architectures; and,
- Resilient Deployment inside Contested Military Environments.

In that sense, the Pentagon is not just rewriting AI contracts. It is rewriting the market for defense-grade AI governance.

## Section VI: Bottom Line

The most important takeaway is brutally simple: the Pentagon has decided that no vendor holds the trigger. Not legally. Not operationally. Not doctrinally. Commercial AI will absolutely

shape future planning, ISR exploitation, command support, and force-wide decision advantage. But the Department is moving aggressively to ensure that this happens inside a framework that protects human accountability, national command authority, mission assurance, and operational continuity.

That is the real story. Not a Silicon Valley feud. Not a press-cycle fight. A foundational rewrite of how military AI will be governed, contracted, contained, and deployed in the age of operational-speed autonomy. The companies that understand this now will shape the future force. The ones that do not will discover, usually too late, that defense AI is no longer just about capability. It is about who can be trusted inside the chain of mission execution.

---

## ✨Call to Action✨:

Subscribe to my LinkedIn Newsletter for more Future Trends and Insights on/in Regulatory Compliance, A.I. (Artificial Intelligence) Governance/Ethics, Internal/External Audit, TPRM (Third-Party Risk Management), Fintech Innovations, and "all things GRC" (Governance, Risk Management, and Compliance). Let's navigate the future together.

---

## 🔍 About the Author:

W.B. Carrington, MBA, LSSMBB, CTPRA, GRCA, GRCP, is a seasoned Artificial Intelligence (A.I.), National Defense and National Security Strategist and Subject Matter Expert (SME) in Governance, Risk, and Compliance (GRC), A.I. Governance/Ethics, Regulatory Compliance, Data Governance, Comprehensive/Integrated Risk Management, and Project/Program Management.

As the Managing Partner at G1VE | Technologies and a Strategic Board Advisor to many multi-domain, cross-industry organizations across both the Public Sector (I.e., Fortune 500 and Fortune 1000 companies) and Private Sector entities, W.B. leverages his extensive experience and deep subject matter expertise to drive innovative solutions that enhance organizational resilience and operational integrity. Connect with W.B. Carrington, MBA, LSSMBB, CTPRA, GRCA, GRCP, on LinkedIn to stay updated on the latest insights and trends in these critical areas.