



Pentagon-Anthropic Dispute over Autonomous Weapon Systems: Potential Issues for Congress

March 13, 2026

On February 27, 2026, President Donald J. Trump [directed](#) federal agencies to “IMMEDIATELY CEASE all use of [American AI company] Anthropic’s technology.” Secretary of Defense Pete Hegseth (who is now using “Secretary of War” as a “secondary title” under [Executive Order \(EO\) 14347](#) dated September 5, 2025) subsequently [directed](#) the Department of Defense (DOD, now using “Department of War” as a secondary designation under [EO 14347](#)) to designate [Anthropic](#) a supply-chain risk to national security; barred defense contractors, suppliers, and partners from working with Anthropic; and described an up-to-six-month period of transition away from Anthropic products. This designation follows a [reportedly](#) months-long dispute between DOD and Anthropic over DOD use of Anthropic products, including the company’s [generative AI](#) model, Claude. On March 9, Anthropic filed a [civil complaint](#) in the U.S. District Court for the Northern District of California and a [petition for review](#) in the U.S. Court of Appeals for the D.C. Circuit regarding these directives.

Some lawmakers have [called for](#) a resolution to the disagreement and for Congress to [act](#) to set rules for the department’s use of AI and/or autonomous weapon systems.

Background

In July 2025, DOD [announced](#) that it had awarded contracts to Anthropic, Google, OpenAI, and xAI for up to \$200 million each “to accelerate Department of Defense (DoD) adoption of advanced AI capabilities to address critical national security challenges.” Although DOD has not publicly outlined the full range of use cases for these companies’ AI models, Anthropic has stated that Claude “is reportedly the Department’s most widely deployed and used [frontier AI](#) model, [and the only frontier AI model on the Department’s classified systems](#).” Anthropic has further [stated](#) its models are used “across the Department of War and other national security agencies for mission-critical applications, such as intelligence analysis, modeling and simulation, operational planning, cyber operations, and more.” [Reports](#) additionally indicate that Claude was used in the January 2026 [operation](#) to capture Venezuelan President Nicolás Maduro, [reportedly prompting](#) inquiries from Anthropic—to the Pentagon—about how its models were

Congressional Research Service

<https://crsreports.congress.gov>

IN12669

being deployed. Anthropic’s [usage policy](#) prohibits, for example, the use of its models to incite violence or to develop or design weapons.

[According to](#) reporting, these inquiries generated concerns within DOD that Anthropic might not approve of certain use cases and, therefore, might attempt to limit DOD use of its models. As a result, the Pentagon [reportedly](#) requested that Anthropic—and other AI companies—allow use of AI models for “all lawful purposes.” While Anthropic was [reportedly](#) “willing to adapt its usage policies for the Pentagon,” the company was, given its [assessment](#) of “what today’s technology can safely and reliably do,” [unwilling](#) to allow two use cases: mass domestic surveillance and fully autonomous weapon systems. In explaining his decision to deny the Pentagon’s [request](#) for “full, unrestricted access to Anthropic’s models,” Anthropic CEO Dario Amodei [stated](#) that autonomous weapon systems “may prove critical for our national defense. But today, frontier AI systems are simply not reliable enough to power fully autonomous weapons.”

DOD is not publicly known to be using Claude—or any other frontier AI model—within autonomous weapon systems. [DOD Directive \(DODD\) 3000.09](#), “Autonomy in Weapon Systems,” outlines the approval process for developing and deploying autonomous weapon systems and identifies requirements for their use.

What Are Autonomous Weapon Systems?

DODD 3000.09 defines autonomous weapon systems as “weapon system[s] that, once activated, can select and engage targets without further intervention by [a human] operator.” This concept of autonomy is also known as *human out of the loop* or *full autonomy*. The directive contrasts such systems with human-supervised, or *human on the loop*, autonomous weapon systems, in which operators have the ability to monitor and halt a weapon’s target engagement. Another category is semi-autonomous, or *human in the loop*, weapon systems that “only engage individual targets or specific target groups that have been selected by [a human] operator.”

DODD 3000.09 requires that all systems, including autonomous weapon systems, be designed to “allow commanders and operators to exercise appropriate levels of human judgment over the use of force.” Such judgment does not require manual human “control” of the weapon system, but rather broader human involvement in decisions about how, when, where, and why the weapon will be employed (i.e., a human must assess the operational environment and decide to deploy the weapon, which can then operate autonomously). This involvement includes a human determination that the weapon will be used “with appropriate care and in accordance with the law of war, applicable treaties, weapon system safety rules, and applicable rules of engagement.” The requirement for “human judgment over the use of force” does not mean that such systems are operating with a *human in the loop*.

For additional information about U.S. policy concerning autonomous weapon systems, see CRS In Focus IF11150, *Defense Primer: U.S. Policy on Lethal Autonomous Weapon Systems*.

Related Legislation and Issues for Congress

The department updated DODD 3000.09 in January 2023, and later that year, Congress passed the National Defense Authorization Act for Fiscal Year 2024 (NDAA; P.L. 118-31), which requires in Section 251 that the Secretary of Defense notify the defense committees of any changes to DODD 3000.09 within 30 days. The Secretary is directed to provide a description of the modification and an explanation of the reasons for the modification. Section 1066 of the FY2025 NDAA (P.L. 118-159) additionally requires the Secretary to submit to the committees, annually through December 31, 2029, a “comprehensive report on the approval and deployment of lethal autonomous weapon systems by the United States” through December 31, 2029. Congress has not legislated on the department’s use of AI models or their reliability.

Should Congress decide that more oversight is needed, it may codify the requirements of DODD 3000.09 or consider additional notification requirements for DOD's use of autonomous weapon systems or AI models. Congress may also restrict funds for the development and/or use of autonomous weapon systems, or for certain use cases of AI models by DOD, should Congress deem such uses as posing an unacceptable level of risk at the current stage of technological development.

Author Information

Kelley M. Saylor
Specialist in Advanced Technology and Global Security

Disclaimer

This document was prepared by the Congressional Research Service (CRS). CRS serves as nonpartisan shared staff to congressional committees and Members of Congress. It operates solely at the behest of and under the direction of Congress. Information in a CRS Report should not be relied upon for purposes other than public understanding of information that has been provided by CRS to Members of Congress in connection with CRS's institutional role. CRS Reports, as a work of the United States Government, are not subject to copyright protection in the United States. Any CRS Report may be reproduced and distributed in its entirety without permission from CRS. However, as a CRS Report may include copyrighted images or material from a third party, you may need to obtain the permission of the copyright holder if you wish to copy or otherwise use copyrighted material.