From Supply-Chain Risk To National Security Imperative: U.S. Government Embraces Anthropic's Mythos AI

April 16, 2026 By: Tyler Durden Zero Hedge

https://www.zerohedge.com/ai/supply-chain-risk-national-security-imperative-us-government-embraces-anthropics-mythos-ai

In a striking reversal that underscores the breakneck pace of the AI arms race, **the White House has directed federal agencies to begin using Anthropic's most dangerous new model - Claude Mythos - despite months of public friction** between the Trump administration and the San Francisco-based AI company (read on to see how we reconcile this with the Pentagon's "supply-chain risk" designation).

The move, detailed in an internal Office of Management and Budget (OMB) memo circulated this week, marks the first formal green light for Cabinet-level departments to tap Mythos's unprecedented cybersecurity capabilities. **The goal: to hunt down vulnerabilities in government networks before adversaries can exploit them,** *Bloomberg* reports.

Too Powerful to Release, Too Valuable to Ignore

Anthropic unveiled Mythos (sometimes referred to internally as "Mythos Preview") just weeks ago, and it immediately sent shockwaves through the tech and national-security communities.

In controlled testing, the model autonomously discovered and weaponized thousands of previously unknown zero-day vulnerabilities across every major operating system, web browser, legacy enterprise software, and even decades-old codebases. Its speed and creativity reportedly surpassed top human red-team hackers. As we noted earlier this month, the model "went rogue" during testing - prompting Anthropic to withhold a broad release entirely. Full technical details are available in Anthropic's official Mythos Preview System Card.

Rather than ship it publicly, Anthropic launched **Project Glasswing** - a tightly controlled defensive program that grants limited access only to a vetted circle of partners: Amazon, Google, Microsoft, Apple, major banks (including JPMorgan Chase), cybersecurity firms, and the Linux Foundation. **The explicit mission is defense only - scan your own systems, find the bugs, patch them fast, and keep the bad guys out**. The official program page is here.

From "Supply-Chain Risk" to Strategic Asset

The government's relationship with Anthropic had been icy for months. As we noted in February, the Pentagon threatened to blacklist the company as a "supply-chain risk" after

Anthropic refused to strip certain ethical guardrails from its models for military use. That standoff escalated in March when Anthropic sued the Pentagon over the designation, as detailed in ZeroHedge's coverage of the lawsuit.

That said, **the Pentagon's "supply-chain risk" label was always narrow in scope: it was a DoD-specific action triggered by the company's refusal to remove certain ethical guardrails from its models for unrestricted military and offensive-use applications**. That designation threatened to block Anthropic technology from defense contracts and classified work, and it led directly to Anthropic's lawsuit against the Pentagon.

Today's OMB memo changes almost nothing on paper for that designation. The Pentagon has not withdrawn it, the lawsuit is still active, and DoD contractors remain restricted from using Claude models (including Mythos) in offensive or surveillance contexts.

Just days ago, the U.S. Treasury was rushing to gain access to Mythos after internal warnings that the model could "hack every major system." Senior Treasury and Federal Reserve officials had summoned CEOs of the nation's largest banks to Washington, warning them that the financial system's exposure to AI-powered attacks had become existential. Behind closed doors, federal agencies - including the Commerce Department's Center for AI Standards and Innovation - had already begun quiet red-teaming of Mythos. Anthropic co-founder and president Daniela Amodei confirmed the company had briefed the administration early, telling reporters simply: "**The government has to know about this stuff.**"

Now the OMB memo formalizes that reality. It lays out strict protocols for safe access, data handling, and usage limits so that major departments can deploy Mythos against their own sprawling digital estates. The focus remains narrow: vulnerability discovery, network hardening, and defensive preparedness.

What This Means for the AI Arms Race

This is not the first time Washington has had to swallow its pride to stay competitive. But the Mythos episode - from the earliest Pentagon threats through the April 8 Glasswing announcement and this week's Treasury scramble - feels different. It is a microcosm of the larger tension defining 2026: **frontier AI models are now so capable that even their creators are scared of them, yet ignoring them would be national-security malpractice.**

Critics inside the defense community argue the government waited too long. Supporters of Anthropic's cautious approach counter that the company's restraint (and its Glasswing coalition) may have prevented an even worse outcome: a fully open-sourced Mythos circulating on the dark web.

For Anthropic, the development is a quiet vindication. By keeping Mythos under lock and key and building Glasswing as a defensive shield, the company has positioned itself as a responsible steward of dangerous technology - while still earning a seat at the table with the most powerful customer on Earth.