The White House's capricious controls on Anthropic

An opaque approach to US policy risks hampering AI development

Financial Times Europe
17 giugno 2026

On Friday, the US commerce department issued an order restricting the use of Anthropic's newest models, Fable and Mythos, by non-US nationals, due to national security concerns — prompting the company to suspend its latest tools. There is some irony in the Trump administration, which brags about its hands-off approach to tech, blocking the vocally pro-regulation Anthropic. But the way the move was carried out — giving Anthropic 90 minutes to comply — has raised alarms across the AI industry. Without a clear and consistent approach to policing frontier models, the US risks hampering the development of the world's most crucial technology.

The order to Anthropic reportedly came after Amazon researchers raised alarms about potential vulnerabilities in Fable that might allow guardrails to be bypassed and users to gain information about security flaws in other software. Anthropic said it had not been given specific details of the US government's concerns but insisted that a narrow "jailbreak" did not merit recalling a commercial model deployed to hundreds of millions of people. Other systems, including those developed by rivals such as OpenAI, possessed similar capabilities, it added.

The latest cutting-edge models are powerful and the potential national security threats are real. A proper framework to vet the most advanced technologies is needed. The Trump administration recently issued an executive order creating a voluntary framework for US agencies to gain early access to AI models, which, although watered down from a previous version, represented an important first step towards a more coherent approach.

The US administration's concerns over Fable remain unclear and may have real substance. But the abrupt and opaque move against Anthropic's model, which had been tested by the US commerce department, exemplifies the White House's arbitrary approach towards AI policymaking.

Given Anthropic's disagreements earlier this year and ongoing litigation with the Department of Defense over its move to label the company a supplychain risk, there will also inevitably be suspicions that the government's move is politically motivated. Senior administration officials have insisted the export controls were an entirely separate safety intervention.

Whatever the reality, industry insiders have cautioned that the White House's heavy-handed approach risked casting a pall over the deployment of frontier AI. Dozens of top cyber security executives issued an open letter warning that the White House's actions had "taken the best models away from [cyber] defenders, [and] created market uncertainty" without setting out a justification.

Washington's move has also unsettled US allies. Despite the widening lead opened up by American tech companies, it is likely to intensify pressure in Europe and elsewhere to reduce reliance on US technology and bolster tech sovereignty by developing non-US alternatives. The European Commission warned that security concerns were a "shared challenge" and measures taken should "not be discriminatory against partners".

AI needs intelligent regulation that does not hamper innovation, but safeguards around frontier models and previously unidentified risks are required. Ideally, as Anthropic's CEO Dario Amodei has advocated, they should be handled by an arm's-length body akin to the US Federal Aviation Administration. Whatever the model, however, striking the right balance is vital for the development of the technology not just in the US but the rest of the world.