ISSUE 4 · 2026-06-13 · EXPORT CONTROL AND MODEL DEPENDENCY

The day the frontier went dark

Issue 4: What an export-control shutoff teaches MedTech about model dependency

KEY TAKEAWAYS

- At 5:21pm ET the night before publication, Anthropic received a US government export control directive and within hours disabled Fable 5 and Mythos 5 for every customer: not throttled, not geofenced, off.
- The stated trigger is a national security concern tied to a reported method for bypassing Fable's safeguards; the order names foreign nationals inside or outside the US, including Anthropic's own foreign national employees, so the only way to comply was to pull both models for everyone.
- Both the dedicated cyber model (Mythos, which ships with its own skills and harness like Claude Code) and the general-purpose model (Fable, which needs a discovery pipeline wrapped around it) were pulled, because the capability lives in the reasoning, not the label on the product page.
- The triggering jailbreak amounts to asking the model to read a codebase and fix the flaws in it; Anthropic says it validated the same capability is available from other models including OpenAI's GPT-5.5 and is used every day by defenders.
- ELTON has lived this at small scale for a year, blocked by guardrails, forced back onto older models mid-run, and made to rebuild around capability that quietly moved or disappeared; what changed is that a model hundreds of millions of people touch went dark on a few hours' notice.
- Prediction: because the directive is built around nationality, with foreign nationals the restricted class and US persons not the stated target, the path of least resistance is to restore US persons first, probably through identity verification such as submitting an ID.

> **"The dependency is real, the off switch is real, and it does not belong to you."**

(Opening)

Frames the overnight shutoff: an export control directive took both Fable 5 and Mythos 5 offline for all customers within hours. The key point is that the specialist cyber model and the general-purpose model were pulled together, because the vulnerability-discovery capability lives in the reasoning rather than in the product label, and the triggering jailbreak was essentially 'read a codebase and fix its flaws.'

We have lived this at small scale. Now everyone gets to.

For a year the author's team absorbed guardrail blocks, mid-run backtracks onto older models, and vanished capability as their own private problem. What changed overnight is the scale: a frontier model hundreds of millions of people rely on went dark on a few hours' notice, making the dependency, and the government's off switch, visible to everyone.

A prediction, for the record

Because the directive is written around nationality, the fastest restoration path is US persons first, most likely via identity verification. That would bench non-US-citizen staff, turn account sharing into a fast way to lose access, and leverage the 30-day retention already required on Mythos-class models for detection.

What I'd actually do about it

Treat frontier model access as a critical single-source supplier and add it to the risk register, and keep the harness model-portable so the agents, tools, handoffs, and validators stay yours while the reasoning engine remains swappable. Do not build hard dependencies on one vendor's continued willingness or legal ability to serve you.

- Shutoff timing: 5:21pm ET, with both models disabled for every customer within hours.
- Models affected: Fable 5 and Mythos 5.
- Directive scope: any foreign national inside or outside the US, including Anthropic's own foreign national employees.
- Comparable capability named in another model: OpenAI's GPT-5.5.
- Scale: a model that hundreds of millions of people touch went dark on a few hours' notice.
- Telemetry cited: Anthropic already requires 30-day data retention on Mythos-class models.