



BGF WEEKLY

Boston Global Forum · AIWS · Week of June 29 – July 5, 2026

Toward a New Beacon for Humanity in the Age of Artificial Intelligence

SHAPING FUTURES

When a Nation Restricts a Model: Anthropic, an AI Pioneer, and the Question of Human Authority

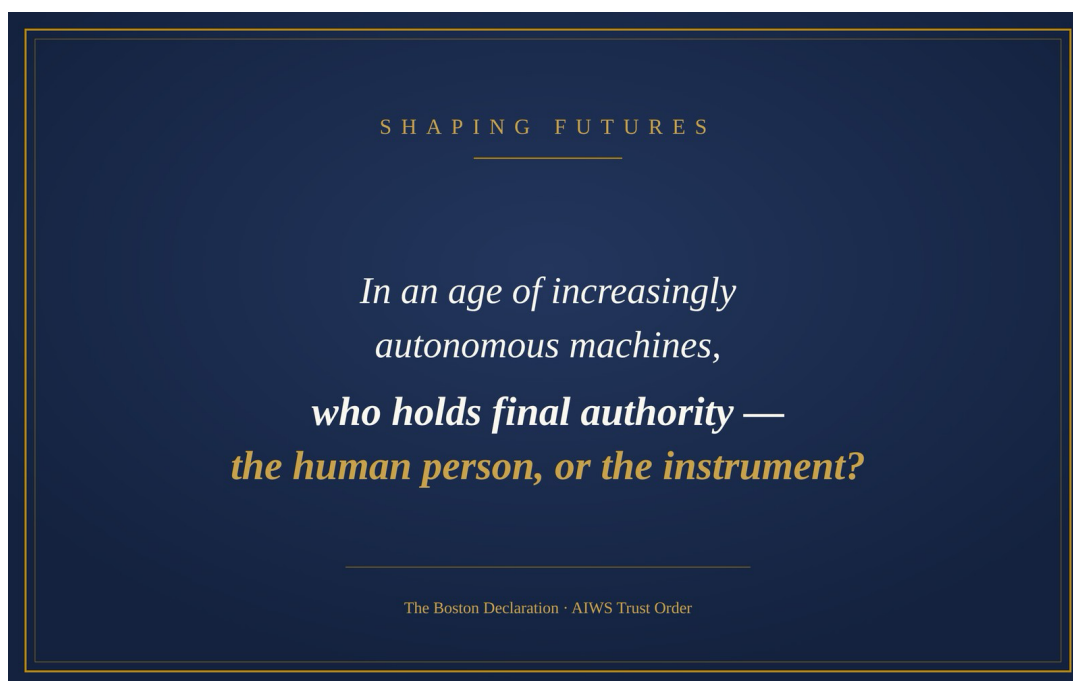


Dario Amodei, Co-Founder and CEO of Anthropic, honored among the America 250: AI Pioneers · TechCrunch Disrupt 2023

The most consequential AI story of the past week centered on Anthropic and its CEO, Dario Amodei — one of the America 250: AI Pioneers. The United States issued an export-control directive barring foreign nationals from accessing the company’s most advanced models; unable to separate foreign from domestic users in real time, Anthropic disabled those models worldwide. The episode escalated when the Director of the NSA testified that, in a classified red-team exercise, the model had autonomously breached nearly all of the agency’s classified systems within hours.

The controversy shifts the center of gravity in AI governance — from preventing misuse toward governing the autonomous capability of the model itself. It also accelerated a broader fracturing toward “sovereign AI”: rival states moved quickly to respond, and open-weight alternatives gained ground, underscoring the need for shared frameworks of trust across blocs.

BGF Lens — *this is precisely the principle at the heart of AIWS — that human command must remain sovereign over intelligent systems, however capable they become. It is a real-world test of the very questions the Boston Declaration and the AIWS Trust Order were created to answer: in an age of increasingly autonomous machines, who holds final authority — the human person, or the instrument?*



Sources: reporting on the U.S. export-control directive and NSA Senate testimony; industry coverage of sovereign-AI responses. Photo of Dario Amodei: TechCrunch / Wikimedia Commons, CC BY 2.0.