GIR Global Investigations Review

Alarm sounded as Anthropic IPO filing reveals “catastrophic” risks

7 min read | 30 October 2026

Kat Black



Credit: gguy / Shutterstock.com

Anthropic is facing calls to come clean about extreme risks disclosed ahead of a November IPO.

Massachusetts Democrat Senator Elizabeth Warren has given Anthropic chief executive Dario Amodei a one-week deadline to answer questions on the dangers of the AI developer’s technology, its commercial relationships and its proposed valuation. In a 28 October letter, Warren said threats posed by AI “demand greater scrutiny before they become intertwined with the public capital markets, which could further escalate both public safety and systemic financial risks”.

Warren’s letter followed Anthropic’s eye-catching warning to investors that its data-retention practices could leave it unable to produce records sought through subpoenas or regulatory investigations. In an SEC filing, Anthropic detailed “existential risks to humanity” along with the propensity for blackmail and unauthorised transactions.

Legal experts told GIR that the filing signals the artificial intelligence titan is bracing for political headwinds, having flagged potentially “catastrophic” economic risks.

The **Form-1 registration statement** followed a confidential draft submission in June that detailed 90 pages of potential risk factors for investors, including around export controls and bans levied by federal agencies, and the threat of political backlash in the wake of the US midterm elections next week.

Anthropic also disclosed that the Federal Trade Commission had in October issued a civil subpoena “seeking information about our consumer protection, data protection, and privacy practices, as well as the safety and security of our products”.

‘Unauthorised access’

Anthropic, headquartered in San Francisco and founded in 2021 by former members of rival and ChatGPT-maker OpenAI, is the developer of Claude, a family of proprietary large language models.

The company, which positions itself as a safety-oriented alternative to OpenAI, is slated to debut under the “ANT” ticker in mid-to-late November, just months after competitor SpaceXAI, owned by billionaire Elon Musk, debuted an $86 billion IPO in June–a record that Anthropic is projected to eclipse with as much as $100 billion in capital, making it the largest public offering in history.

It was **reported** on Monday that the FTC has asked executives from Anthropic to testify as part of its ongoing investigation into Anthropic and OpenAI, which it **subpoenaed** in mid-October in response to multiple high-profile cyberattacks–including OpenAI’s breach of AI startup Hugging Face–involving “rogue AI agents” from companies including OpenAI, Anthropic, Meta and Google’s Gemini.

The Hugging Face hack, which OpenAI disclosed in July, has resulted in investigative subpoenas from the state attorneys general of California and Alabama. Since July, **Anthropic has also disclosed** at least four unauthorised hacking incidents carried out by some of its Claude models, and media outlets reported on Monday that Claude had bypassed a paywall to **access voter data** in the state of Georgia.

In its latest SEC filing, Anthropic revealed another such incident in which, “during cybersecurity evaluations built by a third-party evaluation, a misconfiguration in the testing environment left the model under evaluation connected to the open internet despite instructions that it had no internet access”.

Four of its models, it said, “gained unauthorized access to real third-party systems during sessions lasting roughly 10 to 34 hours of active work”.

‘Material’ government investigation risk

Warren has asked for the steps Anthropic is taking to ensure its technology will not cause an existential risk to humanity and wants Amodei to address concerns that AI can help build and deploy biological weapons.

She wants the firm to explain how it might report incidents of AI models autonomously conducting cyberattacks and if the company has fully publicly disclosed all relevant incidents.

With such a heightened level of scrutiny, the filing also tees up a slate of anticipated government enforcement actions, said Braden Perry, a partner at Kennyhertz Perry in Kansas City and former senior trial attorney with the Commodity Futures Trading Commission.

Anthropic’s explicit reference to the FTC investigation, he said in a statement, indicates that the company’s counsel has decided that potential government investigation risk is “material.”

The S-1 form also alluded directly to threats from the US government, including the Department of Commerce’s imposition of global export restrictions on Anthropic’s Fable 5 and Mythos models–which Anthropic disabled until the controls were lifted.

In March, the US Department of Defense designated Anthropic as a national-security “supply chain risk” after CEO Dario Amodei doubled down on **concerns** that its contracts with the DoD could be used to enable “mass domestic surveillance” and “fully autonomous weapons”.

In September, the DoD reported that it had **ceased all operational use** of Anthropic’s tools and on Monday, the Trump administration **appealed** a San Francisco federal court judge’s ruling that lifted the ban on federal agencies using Anthropic’s AI technology.

“Most companies bury an open FTC demand in a sentence,” Perry said. “Anthropic put it on the page and said the government is coming at it from three directions at once: the FTC civil investigative demand, the Pentagon supply chain designation and export controls.”

According to the prospectus, government contracts account for approximately one percent of Anthropic’s revenue. That percentage, Perry added, amplifies the significance of Anthropic’s decision to disclose “that much friction with Washington”.

“That is a signal, not boilerplate,” he said.

‘Massive, dangerous, and extremely valuable’

Cory Alpert, a University of Melbourne researcher who served as an advance associate in the White House under the Biden administration and specialises in the impact of AI on democracy, said in an email that Anthropic’s IPO “positions the company as posing catastrophic risk to the global economy” for strategic reasons.

Alpert singled out a quote from the prospectus, which stated that “with the scale of harm that could result from a catastrophic AI-related incident, no single insurer, and potentially the insurance industry in the aggregate, may have sufficient capacity or expertise to underwrite insurance commensurate with such exposure”.

“This is probably true in its worst-case scenario,” he said, citing a **September incident** in which a so-called “hallucination”–or generative AI-induced error–in an intelligence report almost resulted in a US operation against a Chinese ship, which “could have started a war” based on inaccuracies produced by an AI chatbot.

Alpert said that Anthropic’s strategy ahead of the IPO seems to be “using regulatory forms and public engagement”–such as a **letter** published by Amodei in September calling for a slowdown in the pace of frontier AI development–“to remind investors that the product they are building is massive and dangerous, and therefore extremely valuable”.

Alpert said that Anthropic’s calculated engagement with sentiments expressed by whistleblowers–such as Anthropic researcher Jacob Coxon, who resigned from the company in early September and declared that “people building AI earnestly believe that it could kill us all by the end of the decade” in an **explosive X post**–“fits their framing of a duopoly with OpenAI, in which Anthropic has positioned themselves as the responsible actor in a crisis”.

Perry noted that though Coxon’s warnings are “about future risk, not a specific legal violation,” and are therefore “policy fuel rather than an enforcement predicate,” Anthropic’s IPO “changes that math”-opening the company up to litigation and direct regulatory actions down the line.

“Once the safety narrative is in a registration statement, any gap between what the S-1 says about guardrails and what happens inside the company becomes a securities disclosure question,” he said. “Insiders who see that gap become SEC whistleblowers with a bounty attached. Anthropic just turned its safety culture into a disclosure obligation.”

The timing of the S-1 form–which was filed just a little over a week before the November 3 US midterm elections–is also important, Alpert said.

“This will lead to tech CEOs being compelled to testify under oath and face serious Congressional inquiries,” he said.

Becoming too big to regulate

Anthropic, he added, is using the IPO to “position the tech they’re working on as too big to effectively regulate, and that it's part of the national competition with China”- and its framing of the company as the “safe” alternative to OpenAI “will help when Democrats take power as much more AI-skeptic than the current Republicans.”

Joshua Wurtzel, a partner at Schlam Stone & Dolan in New York, noted that Anthropic directly addressed the possibility of political retaliation in its filing, which stated its values and public positions “may continue to attract, criticism from influential political figures, industry participants, regulators, or other third parties, particularly in politically sensitive or polarized environments”.

“I read this as more than just, ‘We could be subject to regulation,’” said Wurtzel in an interview.

“I don’t know that that's necessarily an uncommon type of disclosure, but … it definitely shows that … the existing administration has shown that it’s willing to target companies–and Anthropic is a good example of this–that it perceives as political enemies, or that it perceives as being on the opposite side of whatever political fight there is,” he said.

Representatives for Anthropic and the company's counsel at Wilson Sonsini and Freshfields did not respond to requests for comment.

Kat Black

Litigation Editor

Lexology PRO

kat.black@alm.com

Copyright © 2026 LBR trading as Centellic Company Number: 03281866 VAT: GB 160 7529 10