

# Anthropic's New Mythos A.I. Model Sets Off Global Alarms

Paul Mozur, Adam Satariano



Credit...Pablo Delcan

Mythos has triggered emergency responses from central banks and intelligence agencies globally, as Anthropic decides who has access to the powerful model.

Credit...Pablo Delcan

Listen · 8:56 min

By [Paul Mozur](#) and [Adam Satariano](#)

Paul Mozur, based in Taipei, Taiwan, and Adam Satariano, in London, cover global technology issues.

April 22, 2026

When [Anthropic told the world](#) this month that it had built an artificial intelligence model so powerful that it was too dangerous to release widely, the company named 11 organizations as partners to help mount a defense.

All were from the United States.

Within two weeks, the model, called Mythos, had set off a global scramble unlike anything yet seen in the A.I. era. Mythos, which Anthropic has said is uncannily capable of finding and exploiting hidden flaws in the software that runs

[the world's banks](#), power grids and governments, had become a geopolitical chip — and a U.S. company held it.

World leaders have struggled to figure out the scale of the security risks and how to fix them, with [Anthropic](#) sharing Mythos with only Britain outside the United States. The Bank of England governor warned publicly that Anthropic may have found a way to “crack the whole cyber-risk world open.” The European Central Bank began quietly questioning banks about their defenses. Canada’s finance minister compared the threat to the closure of the Strait of Hormuz.

For U.S. rivals like China and Russia, Mythos underscored the security consequences of falling behind in the A.I. race. One Russian pro-Kremlin outlet called the model “worse than a nuclear bomb.”

The responses illustrated a reality that A.I. researchers have long warned about mostly in theoretical terms: Whoever leads in building the most powerful A.I. models will gain outsized geopolitical advantages. Major A.I. breakthroughs are beginning to function less like product launches and more like weapons tests, and most nations want to understand how the technologies work and what protections are needed.

As foundational A.I. “models become more consequential, access becomes more geopolitical,” said Eduardo Levy Yeyati, a former chief economist at the Central Bank of Argentina and a regional adviser on growth and A.I. at the Inter-American Development Bank. “I would take this episode as a policy wake-up call. Governments can no longer ignore the issue.”

Even the U.S. government, which has been [embroiled in a clash](#) with Anthropic over the use of A.I. in warfare, has taken notice of Mythos. On Friday, Dario Amodei, Anthropic’s chief executive, [met with White House officials](#) after some in the Trump administration noted the potential for the new model to wreak havoc on computer systems.

Image



*Anthropic’s chief executive, Dario Amodei, met with White House officials on Friday in a meeting that was described as “productive.” Credit...Karsten Moran for The New York Times*

Anthropic, which is based in San Francisco, told The New York Times that it was keeping access to Mythos small because of safety and security concerns. It has focused on sharing the model with more than 40 organizations that provide technology used in maintaining critical global infrastructure like the internet or electricity grids. Anthropic named 11 of the organizations, including Amazon, Apple and Microsoft, that pledged to help develop security fixes for vulnerabilities identified by the model.

The company said that it had no immediate timeline for widely expanding access, but that it would work with the U.S. government and industry partners to determine next steps. It said that it had been bombarded by calls from governments, companies and other organizations seeking access and information, but that these organizations could have varying levels of expertise to safely evaluate such a powerful A.I. model.

Anthropic added that it expected other groups to release A.I. models with similar cyber capabilities more widely within at least 18 months, giving organizations limited time to make the necessary security fixes.

On Tuesday, Anthropic said it was investigating a report that unauthorized users gained access to a version of Mythos.

The scramble over Mythos comes at a moment of minimal international cooperation on A.I. Governments are viewing one another with suspicion as corporations race to outpace rivals. There is no equivalent of the Nuclear Nonproliferation Treaty, no shared inspections and no agreed-upon rules for how to handle something like Mythos.

When Anthropic announced the model, many experts praised the company's caution in limiting who gets to try the model, but expressed concerns about the lack of international coordination to deal with the risk.

Britain was the only other nation to gain access. Its A.I. Security Institute, a government-backed organization, tested Mythos and [published](#) an independent evaluation last week, confirming that it could carry out complex cyberattacks that no previous A.I. model had completed.

“This represents a step up in A.I. cyber capabilities,” Kanishka Narayan, Britain's A.I. minister, said last week on social media, saying the country was taking steps to protect “critical national infrastructure.”

Others got less information. The European Commission, the executive branch of the 27-nation European Union, has met with Anthropic at least three times since the Mythos release, an E.U. official said. But the company has not provided access to the model because the two sides have not agreed on how to share it with the commission, the official said.

Image



*The European Commission, the executive branch of the European Union, has met with Anthropic at least three times since the Mythos release. Credit...Frederick Florin/Agence France-Presse — Getty Images*

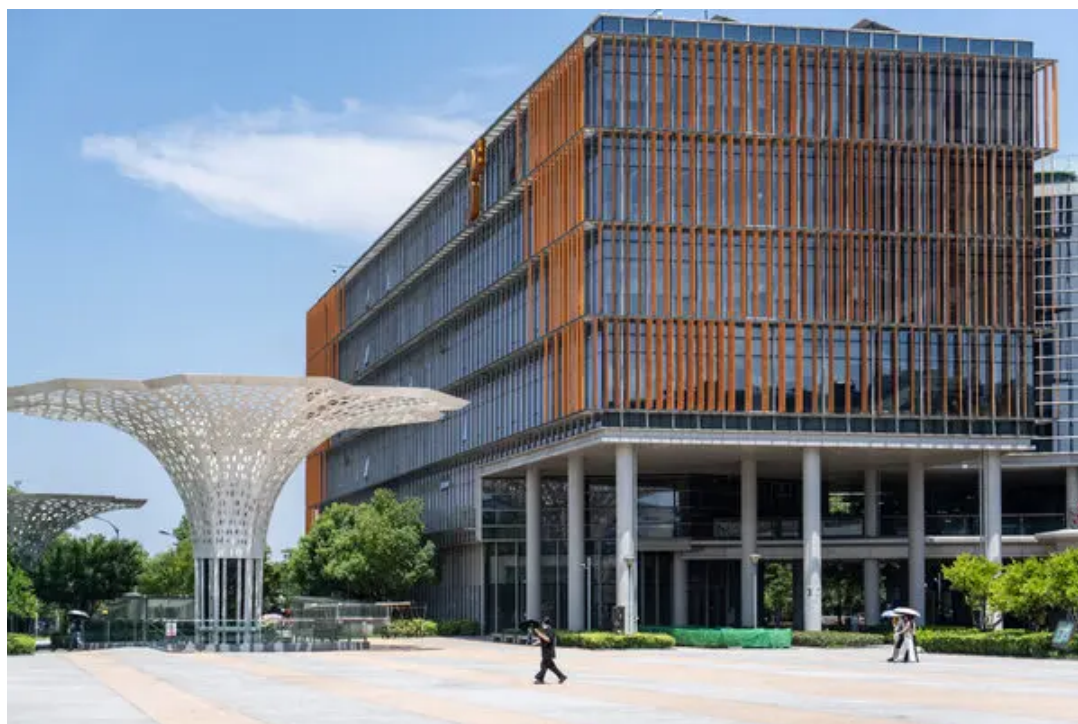
In a statement, the commission said it was “assessing possible implications” of Mythos, which “exhibits unprecedented cyber capabilities.”

Claudia Plattner, the president of Germany’s cybersecurity agency, known as B.S.I., said it had not received access to Mythos, but she met with Anthropic employees in San Francisco recently for “meaningful insight” into how it works. The capabilities point to “a paradigm change in the nature of cyber threats,” Ms. Plattner said in a statement.

Among U.S. rivals, the response has been more muted. Despite Anthropic’s [recent clash](#) with the Trump administration, Mr. Amodei has made clear that A.I. should be used to defend the United States and other democracies and defeat autocratic adversaries.

Neither Beijing nor Moscow has made a major public statement on Mythos. Inside China, researchers and the broader A.I. community have been watching closely, according to analysts studying the country’s tech community. Many of the country’s banks, energy companies and government agencies run on the same software in which Mythos found vulnerabilities — but for now, they have no seat at the table.

Image



*Alibaba Innovation Park in Hangzhou, China. Inside China, researchers and the broader A.I. community have been watching Mythos closely. Credit...Qilai Shen for The New York Times*

“For China I think this is the second wake-up call after ChatGPT,” said Matt Sheehan, a senior fellow at the Carnegie Endowment for International Peace. He added that a U.S. policy to prevent China from obtaining the most sophisticated semiconductors for building advanced A.I. systems was helping to extend the U.S. lead.

Some A.I. researchers in China have privately expressed concern that the country could fall further behind, missing out on advantages that come with building a foundational model first, said Jeffrey Ding, a professor of political science at George Washington University.

Liu Pengyu, a spokesman for the Chinese Embassy in Washington, said China was not familiar with the specifics of Mythos but supported a peaceful, secure and open cyberspace.

Mythos is the latest sign of a growing [global A.I. divide](#). Nations without powerful computing infrastructure and A.I. models risk being left dependent on companies like Anthropic, Google and OpenAI while having little sway over how their products are designed and safeguarded, Mr. Yeyati said.

“The idea that access to frontier A.I. is something a company can unilaterally restrict, using criteria that are opaque and

unappealable, should be a real concern,” he said.

[Paul Mozur](#) is the global technology correspondent for The Times, based in Taipei. Previously he wrote about technology and politics in Asia from Hong Kong, Shanghai and Seoul.

[Adam Satariano](#) is a technology correspondent for The Times, based in London.

A version of this article appears in print on April 24, 2026, Section B, Page 1 of the New York edition with the headline: New A.I. Model Spurs a Scramble. [Order Reprints](#) | [Today's Paper](#) | [Subscribe](#)

Advertisement

[SKIP ADVERTISEMENT](#)