What is Anthropic's Claude Mythos model?

Is Mythos really that capable? Why is this LLM not being made available to the public? What is Project Glasswing?

Published – April 12, 2026 03:58 am IST

AROON DEEP

Image used for representational purposes. File | Photo Credit: Reuters

The story so far:

On April 7, Anthropic, the AI company behind the coding and productivity-focused Large Language Model (LLM) family Claude, announced Mythos. This is its most powerful model yet, capable of finding bugs in old software that have not been flagged by humans so far. Anthropic said that it would not release the LLM widely, but only to a consortium of over 40 companies, which will use it to scan decades-old code to find software vulnerabilities not detected by humans yet.

What is Claude and why is it a notable product in AI circles?

Claude is an LLM developed by San Francisco-based Anthropic, like OpenAI's ChatGPT and Google's Gemini. However, its reputation for quality outputs in fields such as coding have lent it a reputation unlike any other LLM on the market. Claude runs through a command-line interface as well as a suite of apps for different platforms, published by Anthropic.

In increasing order of sophistication, Anthropic's other LLMs are Haiku, Sonnet, and Opus. All three have been noted for their performance and reasoning capabilities, which are critical to composing code and performing tasks agentically. AI firms often don't open-source their models, restricting access by rationing access to use and employing subscription- or usage-based pricing for users.

In spite of generally significant pricing for both usage-based pricing and subscriptions, most AI companies are not profitable, and are staking huge amounts of operational expenditure in the hopes of coming out on top in the AI race. Anthropic is not an outlier, since it does not make profits, but its usage limits and pricing are the subject of frequent complaints among users.

How is Mythos different from Claude's other models?

Cybersecurity has been an unintended side effect of these models' coding prowess. Opus was able to find multiple bugs in highly scrutinised pieces of open-source software, which are used in both private and public IT systems heavily. Human reviewers frequently find bugs, or security vulnerabilities, and they are "patched" so that attackers cannot exploit them to remotely shut down or gain access to computer systems.

Opus found bugs that humans missed, one Anthropic executive said at a talk in March, leading to worries within the company that its technology could be used by hackers to exploit, rather than find and patch, vulnerabilities.

Mythos has already been able to find "hundreds" of "severe" security vulnerabilities. Anthropic announced Project Glasswing, a defensive cybersecurity initiative, in partnership with Microsoft, Apple, Cisco, and other companies whose products and services form the backbone of several other companies and products across cyberspace.

Is Mythos really that capable? What is the problem if everyone has access?

It is not possible to know exactly how capable Mythos is since only a select group has access to it. But the fact that Anthropic has been able to announce specific vulnerabilities it has been able to co-develop patches for in established pieces of open source software indicates that its value as a cybersecurity tool is too significant for large IT and software firms to ignore.

The issue with making Mythos available generally is a transitional one. While Mythos is arguably the first to develop a model that can identify vulnerabilities, it is likely that other models with these capabilities will appear eventually. Anthropic's logic with Project Glasswing is that if the companies and individuals developing these foundational systems have access, they can get a head start in plugging vulnerabilities before attackers gain access to Mythos-class models and start attempting cyberattacks with these capabilities.

What are the implications for India? Is the government doing anything?

The Indian IT industry relies on a range of foreign platforms and software, and often builds its own bespoke software solutions. If Project Glasswing finds a wide range of bugs before Mythos-class models proliferate, Indian companies would benefit by patching all the software they use in time. However, their own software may be vulnerable to sophisticated attackers. No Indian IT firm has been publicly listed as a Project Glasswing partner as yet. The Data Security Council of India under Nasscom has been holding meetings on Mythos over the last week, its CEO Vinayak Godse told *The Hindu*. The IT Ministry and its subordinate Indian Computer Emergency Response Team have been studying the implications of Mythos, a senior government official told *The Hindu*.

Published – April 12, 2026 03:58 am IST

In Case You Missed It