



Sanskriti IAS

25th April 2026



IMPORTANT

EDITORIAL HIGHLIGHTS

DELHI CENTRE:
636, Mukherjee Nagar
New Delhi-110009

PRAYAGRAJ CENTRE:
1/1/8A, Stanley Rd,
Maharana Pratap Chauraha,
Civil Lines, Prayagraj, UP - 211002

 **9555-124-124**

 **sanskritiias.com**

GS 2: INDIAN POLITY

INDIAN EXPRESS PAGE : 1

WELCOMED BY NABIN INTO BJP FOLD; SHOCK WAVES IN PUNJAB WHERE POLLS ARE DUE EARLY NEXT YEAR

In major setback to AAP, seven of its 10 Rajya Sabha MPs cross over to BJP

AAP calls it 'betrayal', accuses BJP of another 'Operation Lotus'

Jatin Anand & Liz Mathew
New Delhi, April 24

IN A huge setback to the AAP, seven of its 10 Rajya Sabha members, constituting two-thirds of the party strength in the Upper House, crossed over to the BJP Friday, creating ripples in political circles, especially in Punjab where AAP's rule will be tested in Assembly elections early next year.

Addressing a press conference in New Delhi, AAP MP Raghav Chadha, removed as the party deputy leader in Rajya Sabha three weeks ago, said he, along with six party members of the House, had "merged with the BJP".

The AAP had 10 MPs in Rajya Sabha — 7 from Punjab, 3 from Delhi.

Flanked by MPs Ashok Mittal and Sandeep Pathak, Chadha announced their exit from the AAP along with Rajinder Gupta, Vikramjeet Singh Sahney, Swati Maliwal and Harbhajan Singh.

"They have already signed, and this morning we submitted all the required documentation, including signed letters and other formal paperwork, to the Chairman of the Rajya Sabha," he said.

Following the press conference, Chadha and his colleagues headed to the BJP headquarters where they were

»CONTINUED ON PAGE 2



Raghav Chadha, Sandeep Pathak and Ashok Mittal with BJP president Nitin Nabin and party leaders Anil Baluni (extreme left), Tarun Chugh and Arun Singh (extreme right) at the BJP headquarters in New Delhi, Friday. (PI)

Chadha's removal, Mittal on ED radar: AAP points to early signs

Jatin Anand & Deeptiman Tiwary
New Delhi, April 24

STUNNED BY the exit of seven of their 10 Rajya Sabha members, AAP insiders said Friday they were aware of a storm brewing for the last "week to ten days" following the party decision to remove Raghav Chadha, one of its most visible faces, as the AAP deputy leader in the Upper House, literally showing him the door.

According to party insiders, AAP national general secretary (organisation) Sandeep Pathak, a close aide of party chief Arvind Kejriwal, had called on the former Delhi Chief Minister at the latter's residence in the Capital when he was in town during the special session of Parliament last week.

While details related to Pathak's meeting with Kejriwal remained unclear, sources said he was likely to have conveyed that he no longer enjoyed much support from the party's organisational ranks in Punjab where he played a significant role in building the AAP.

"We were more than aware of some trouble brewing over the past week to ten days; party workers were aware of several meetings between both Raghav (Chadha) and Swati (Maliwal) with senior BJP leaders in Delhi," a party source claimed.

That something was afoot, insiders said, became clear following the Enforcement Directorate's searches at Rajya Sabha member Ashok Mittal's

»CONTINUED ON PAGE 2

"The AAP which I nurtured with my blood and sweat... has strayed from its principles, values and core morals."

RAGHAV CHADHA, AHEAD OF JOINING BJP

"Game of dirty politics... to stall the good work of the Bhagwant Mann government (in Punjab) has begun on a large scale."

SANJAY SINGH, AAP RAJYA SABHA MP

AAP in Parliament

RAJYA SABHA: 10

From Punjab: • Ashok Kumar Mittal • Harbhajan Singh • Raghav Chadha • Rajinder Gupta • Sandeep Kumar Pathak • Sant Balbir Singh • Vikramjeet Singh Sahney

From Delhi: • Narain Dass Gupta • Sanjay Singh • Swati Maliwal

LOK SABHA: 3

• Gurmeet Singh Meht Hayer • Malvinder Singh Kang • Raj Kumar Chabbewal All from Punjab

Faultlines surface in AAP Punjab unit, raise fears of wider churn

Kanchan Vasdev
Chandigarh, April 24

THE DEFECTION of seven Aam Aadmi Party (AAP) Rajya Sabha MPs is not merely a setback in Delhi, but could have an effect in Punjab, where the party's government faces Assembly elections in less than a year.

What has the AAP leadership concerned is the deep organisational roles that Raghav

Chadha and Sandeep Pathak had in Punjab and the influence they wielded all these years right from the selection of candidates ahead of the 2022 polls. Nearly 50% of AAP MLAs owe their candidature to either of the two.

In other words, the split in the party's ranks in Rajya Sabha also brings to the fore long-standing tensions

within the state unit. This would test AAP's ability to retain cohesion in a unit that expanded rapidly and is controlled by the central leadership.

E.

Chief Minister Bhagwant Mann slammed the rebel MPs as "gad-daar (traitors) of Punjab", and said they left the party to "save their own skin". Just

»CONTINUED ON PAGE 2

INSIDE PAGES 4, 17
THE BJP'S NEW SEVEN

EXPLAINED: WILL THEY BE DISQUALIFIED?

GS 2: INDIAN POLITY

INDIAN EXPRESS PAGE : 10

Mamata Banerjee's economics doesn't work. Will her politics swing the election?

I HAVE followed and written about elections for the last 30 years. I have had my glorious successes (Lalu Yadav's win in 2015, Narendra Modi's win in 2019) and inglorious failures (Hillary Clinton's loss in 2016 and the BJP's "loss" in 2024). But never before has a forecast so deeply conflicted me.

This week I toured West Bengal (along with other self-confessed election junkies led by Ruchir Sharma), and my most enjoyable moment was a visit to Presidency College, now Presidency University. But that is getting ahead of the story. Throughout my career, I have put a larger emphasis on economic variables for forecasting. Term it the consistency of little minds or any such rejection phrase, but the fact remains that welfare improvement (proxied by state GDP per capita) remains by far the best predictor of winning elections.

As far as economic fortunes go, no one deserves the fate of West Bengal. The culturally rich state was also rich in wealth and industry. Its per capita income (relative to all India) was 113.8 in 1950, which slipped to 90.3 per cent in 1976, the year before the Marxist revolution — an annual decline of 0.9 per cent per annum. Relative income in 2010 was 80.8 per cent; a CPM degradation at -0.3 per cent per annum. However, the worst came with Mamata Banerjee's arrival — the decline in the last 13 years (2011-2024) accelerated to -1 per cent per annum. Today, per capita income in West Bengal is just 70.3 per

cent of the average big-state income. On another occasion, we will compare this decline to other worst-performing states in India and countries in the world. For the moment, note that this macro income performance is the worst India has seen for any state.

But first, the phrase of the century, in India and across the world, *nari shakti*. The overwhelming evidence is that the *nari shakti* vote in West Bengal was particularly ill-timed and ill-thought. Why? Because the BJP is fighting Banerjee, the undisputed king and queen of *nari shakti*. Just look at the electoral evidence. This is an area where Banerjee has done particularly well. In the 2021 Assembly elections, the BJP fielded 57 women candidates, and only won five (8.8 per cent) won. Banerjee's party — the All India Trinamool Congress (AITC) — fielded seven fewer female candidates than the BJP, but AITC women won 80 per cent (40) of the contested seats.

Preliminary analysis of all election data shows that Trinamool consistently leads all major parties in women's representation. It is a moot question whether *nari shakti* among politicians translates into higher *nari shakti* among the people. It likely does. Parenthetically, what Banerjee has done is what is recommended for all political parties in India — allow for greater representation of women in political arenas at all levels of governance. That is



SURJIT S
BHALLA

The reason for my conflicted thinking should be clear — this is a close election. Or was?

A record turn-out helps the challenger, but the discontent induced by SIR could also help the incumbent

true *nari shakti* rather than one forced through yet another reservation.

Improvements in female labour force participation (FLFPR) is normally a healthy sign for an economy. But it can be a double-edged indicator. Poverty also induces a higher participation rate, and as we have seen, West Bengal has consistently become poorer. With that caveat, West Bengal FLFPR rates show a healthy increase. For 15-64 year olds, weekly status, the FLFPR increased from 21.7 per cent in 2011-12 to 35.3 per cent in 2024-5. This increase of 14 percentage points (ppt) is average, not exceptional. In Assam, the increase was 26 ppt, Bihar 18 ppt, Jammu and Kashmir 21 ppt, Odisha and Madhya Pradesh 17 ppt.

If Banerjee wins this election, she will join a very rare group of politicians. A politician winning four-plus consecutive state elections as chief minister has happened in two large states — West Bengal and Odisha — making it one of the rarest achievements in Indian electoral history. No chief minister in other major states like UP, Maharashtra, Bihar, Tamil Nadu, or Madhya Pradesh has achieved this, as those states have consistently seen alternating governments. Incidentally, the statistical record suggests that if Banerjee wins in 2026, she will likely win in 2031.

The reason for my conflicted thinking should be abundantly clear — this is a close election. Or

was? A record turn-out helps the challenger, but the discontent induced by the SIR could also help the incumbent. All things considered, in the beginning of the week I believed Banerjee would win, by heavily reduced margins. Now, with this greater-than-90 per cent turnout (where else besides Russia does one obtain such turnout?) and economic determinants winning, the accelerated record pace of de-development in West Bengal under Banerjee suggests that the BJP will likely hold the edge. But beware, in a close two-way election, small differences in aggregate vote share can mean a larger than thin victory.

If the BJP wins, the question will be, can it win the economy? Private investment has worsened considerably under the BJP (at a record pace). Aggregate investment has been held at 32 per cent due to high public investment. Investments and growth need to increase by a lot for Viksit Bharat to be credible. A further downside to a BJP win is that the already complacent party will become further entrenched in satisfaction. A good upside will be if the BJP learns to introspect on all aspects of its 12-year reign. Along with the Bengal economy, India can also recover to be on the path of Viksit Bharat.

Bhalla is chairperson of the Technical Expert Group for the first official Household Income Survey for India. Views are personal

GS 2: INTERNATIONAL RELATIONS

INDIAN EXPRESS PAGE : 12

WEST ASIA CONFLICT

No BRICS consensus on war, India's Chair summary notes 'deep concern'

Gaza, Lebanon ceasefire, attacks on UNIFIL figure in grouping's discussions

Shubhajit Roy
New Delhi, April 24

WITH OPINION divided within the BRICS grouping on the US-Israel war on Iran, a consensus was elusive and Delhi issued a Chair's summary on Friday and said that "members expressed deep concern on the recent conflict in the Middle East [West Asia] and offered views and assessments on the matter".

The Chair's statement is a declaration by the host and the

chair of any grouping, which is a summary of the statement, but has not been agreed to by all the members. Usually, a joint statement which is agreed to by all the members is issued after any multilateral meeting. This was a practice in the G20 meetings in the run-up to the summit in 2023, when there was no consensus on the Russia-Ukraine war.

India, which holds the chair of BRICS in 2026, said, "The BRICS Deputy Foreign Ministers and Special Envoys on the Middle East and North Africa (MENA) met on 24 April 2026 in New Delhi, and exchanged views on the current situation in the region. Members expressed deep concern on the recent conflict in the Middle East and offered views and assessments on the matter".

The discussions covered the Palestine issue and the Gaza

situation, including provision of humanitarian aid, the role of UNRWA, zero-tolerance approach to terrorism, welcoming the ceasefire in Lebanon, the unacceptability of attacks on UNIFIL, post-conflict reconstruction and rehabilitation in Syria, political settlement in Yemen, stability and development in Iraq, political process in Libya, and addressing the humanitarian crisis in Sudan. They agreed to meet again under China's Chairship in 2027," it said.

In early March, *The Indian Express* had reported that on March 13 Tehran reached out to Delhi to take the lead in issuing a statement on behalf of the BRICS, currently chaired by India, condemning the US and Israeli strikes against Iran in the previous two weeks.

This put Delhi in a diplomatic pickle since it has stayed

away from taking sides in the ongoing conflict.

The BRICS grouping has Brazil, Russia, India, China and South Africa (the original five), along with Egypt, Ethiopia, Iran, Saudi Arabia, the UAE and Indonesia. Saudi Arabia and the UAE, which host US bases and personnel, have been targeted by Iranian missiles and drones. The fact that all three countries are in the BRICS has made diplomatic navigation difficult for India. India, which holds the rotational presidency of the grouping for 2026, is expected to host the summit of BRICS leaders later this year.

With Iran reaching out to India for a statement by BRICS, Delhi had made it public in mid-March that it has been difficult to forge a consensus on the issue because of involvement of BRICS member coun-

tries in the ongoing conflict.

Without naming any country, the Ministry of External Affairs' official spokesperson Randhir Jaiswal had said, "Some members of the BRICS are directly involved in the current situation in the West Asia region, which has impacted forging a consensus on a common BRICS position on the ongoing conflict. As Chair of BRICS, India has been facilitating discussions among members through the Sherpa channel."

He said that the last virtual BRICS Sherpa meeting was held on March 12. In addition, "Indian leadership has been engaging with leaders of BRICS members in the region. India will continue to engage," the MEA spokesperson said.

India is likely to host the BRICS Foreign ministers' meeting on May 14-15, which Chinese

Foreign minister Wang Yi and Russian Foreign Minister Sergey Lavrov are expected to attend, sources said. Iran's Foreign Minister Seyed Abbas Araghchi may join virtually, depending on the circumstances.

While there are many moving parts as a result of the ongoing conflict and the ceasefire talks between US and Iran, sources said, Delhi is preparing the ground for the meeting of the Foreign ministers.

Senior officials of the BRICS countries are in touch with each other for the Foreign ministers' meeting, which will ultimately lead to the BRICS leaders' summit to be held in India later this year. The timing of the BRICS foreign ministers is significant since India has not been able to forge a consensus among the BRICS grouping on the war in West Asia.

GS 3: ENVIRONMENT

INDIAN EXPRESS PAGE : 12

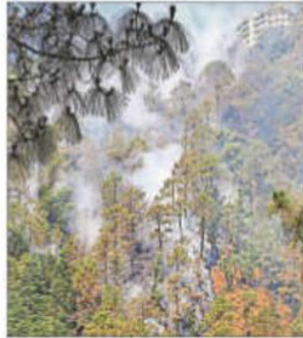
Extreme events could impact 36% of land animal habitats by 2085, says new study

Nikhil Ghanekar
New Delhi, April 24

BY 2085, over a third of wildlife habitats on land could suffer multiple types of climate-driven extreme weather events, such as heatwaves, wildfires, and floods, if warming continues to rise in the second half of the century, according to a new study published in Nature Ecology & Evolution journal.

The study, led by researchers at the Potsdam Institute for Climate Impact Research (PIK), Germany, analysed the impact of climate change in a medium-high emission scenario on nearly 34,000 vertebrate species.

The most widespread impact will be from extreme heatwaves, the study found. By 2050, 74% of animal habitats on land will be exposed to heat-



Extreme climate events can have devastating impacts on terrestrial biodiversity, the study underscored. PT

waves, 16% to wildfire, 8% to droughts, and 3% to floods if warming continues into the latter half of the century, the study said. Biodiversity hotspots such as the Amazon, tropical Africa and Southeast

Asia will see a sharp increase in frequency of these extreme events, the study noted.

However, cutting emissions to net zero could still prevent these impacts, and also limit the animal habitats exposed to extreme events, the study highlighted.

An international team of 18 scientists used climate projections, impact simulations along with data of species distribution to assess exposure risk to 33,936 terrestrial vertebrate species and 794 ecoregions.

"I think climate change, and in particular extreme events, are still really being underestimated when it comes to conservation planning. It's not just going to be a gradual shift of temperature over many years," said lead author Stefanie Heinicke, a postdoctoral researcher at PIK, in a statement.

"There is a lot of difference we can make by cutting emissions as fast as we can from today," Heinicke added.

Extreme climate events, the study underscored, can have devastating impacts on terrestrial biodiversity. The 2019-20 heatwave in Australia, the study said, killed more than 72,000 flying foxes – one of the largest bat species in the world. In the same year, wildfires in the Pantanal (South America) killed an estimated 17 million vertebrates, as per the study.

While one heatwave or wildfire event can devastate animal populations, multiple events occurring one after another compound the impacts. However, climate change, the study said, is intensifying extreme events beyond what many species are likely to adapt to within a short time-frame.

GS 3: SCIENCE AND TECHNOLOGY

INDIAN EXPRESS PAGE : 13

RISK POSED BY MYTHOS

Anthropic in talks with India to safeguard critical infra

Soumyarendra Barik
New Delhi, April 24

ANTHROPIC IS talking to several governments, including that of India, to help safeguard critical infrastructure — banking, energy, telecom — in the face of cybersecurity risks posed by its powerful AI model, Mythos, *The Indian Express* has learnt.

"Anthropic is actively engaging with a broad range of governments, including India. There are ongoing discussions which are very constructive. The conversations are centred around securing critical infrastructure.

There is a belief within Anthropic that allied democracies need access to defence capabilities against powerful AI models," a source close to the company said on condition of anonymity.

The company is also coordinating very closely with the US government on three things — the timing, scope, and the structure of providing access to Mythos to Washington so that they can study the system and prepare defensive measures against its risks.

This particular conversation, centred around offering access to the government, is not currently happening between Anthropic and India, it is understood.

The Indian Express had earlier reported that the government is currently in conversation with Anthropic's senior leadership in the US on the issue, a discussion which was initiated by the Ministry of External Affairs.

Earlier this month, Anthropic unveiled a preview of Mythos, calling it the most powerful model it has ever built. So powerful that unlike typical AI launches, the com-

• PUTTING GUARDRAILS

ANTHROPIC UNVEILED a preview of Mythos earlier this month, calling it the most powerful model it has ever built

UNLIKE TYPICAL AI launches, the company is deliberately keeping Mythos away from the public for now due to its "unprecedented" ability to autonomously identify serious vulnerabilities in widely used software and infrastructure, posing significant cyber security threats



ANTHROPIC PARALLELLY

announced Project Glasswing, which aims to assist companies that hope to use Mythos to step up cyber-defences before it is widely released.

pany is deliberately keeping the technology away from the public for now — the reason being the model's unprecedented ability to autonomously identify serious vulnerabilities in widely used software and infrastructure, posing significant cyber security threats.

Even as it held back commercial deployment citing all these concerns, Anthropic parallelly announced Project Glasswing, which aims to assist companies that hope to use Mythos to step up cyber-defences before it is widely released.

Major software developers—including Apple, Nvidia, the Linux Foundation and CrowdStrike, as well as competitor Google were in this list.

On Thursday, chairing a high-level meeting to assess the risks that Mythos poses to India's financial sector, Union Finance Minister Nirmala Sitharaman told banks to exercise a "high-degree" of vigilance and develop a coordination mechanism to respond to threats emerging from capabilities of the advanced large language model.

In a statement, the Minis-

try of Finance said the meeting took place "to assess the potential impact of emerging threats linked to recent developments in artificial intelligence (AI) models, particularly the possibility of such technologies being misused to weaponise software vulnerabilities".

"...the nature of the emerging threat from the latest AI model is unprecedented and requires a very high degree of vigilance, preparedness and better coordination across financial institutions and banks," the Ministry said, without naming Mythos.

A senior official said a report stating that Mythos, not yet widely released owing to its capabilities, was accessed by unauthorised individuals, had raised significant concerns. Anthropic has said Mythos can outstrip all but the most skilled humans at identifying and exploiting software vulnerabilities, and in tests, the model found critical faults in every widely used operating system and web browser.

FULL REPORT ON
WWW.INDIANEXPRESS.COM

GS 3: SCIENCE AND TECHNOLOGY

INDIAN EXPRESS PAGE : 15

• TECH

The global risks posed by Anthropic's Mythos AI

Anthropic's latest model marks a leap in detecting and exploiting software vulnerabilities. It has spooked policymakers worldwide



Soumya Barik

In 2019, Dario Amodei, then OpenAI's research director, warned that the startup's new large language model was "too dangerous to release" over its potential for generating misleading content. When GPT-2 was eventually released almost a year later, the concerns seemed somewhat overblown.

But fresh warnings by the soft-spoken Amodei, now chief of OpenAI rival Anthropic, about the risks posed by his company's latest AI model — Claude Mythos — appear far more grounded. Mythos was held back from release by Anthropic but now appears to have been leaked.

The model has prompted concerns around the world, and in India. On Thursday, Finance Minister Nirmala Sitharaman chaired a high-level meeting over concerns that it could pose significant risks to India's banking sector. The Indian Express has learnt that the government is currently in conversation with Anthropic's senior leadership in the US on the issue.

Sparking more worry within the government are reports that China has developed its own version of Mythos, called Qihoo 360. This "vulnerability discovery agent" that has already identified nearly 1,000 software flaws. While Mythos is recognised as a breakthrough in automated cyber defence, Chinese state-owned enterprises and specialised firms such as 360 are aggressively advancing similar capabilities.

First, what is Mythos?

Mythos is the latest addition to Anthropic's Claude AI family. Anthropic has unveiled only a limited preview version of the general-purpose language model.

But Mythos represents a major leap in capability over Anthropic's existing AI systems. It has drawn global attention because of its ability to autonomously identify serious vulnerabilities in widely used software and infrastructure at a pace far beyond most human researchers. It is highly capable of fixing security issues (if set to work as a defender) — but also, crucially, at exploiting them (if deployed as a hacker).

Why has this raised concerns?

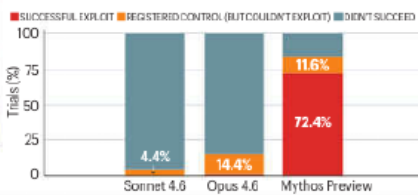
What's spooked policymakers around



Anthropic CEO Dario Amodei. WIKIMEDIA COMMONS

In a different league

During internal testing, Mythos Preview easily outstripped other Claude models in exploiting vulnerabilities — in this case, in Mozilla's Firefox internet browser.



the world is Anthropic's claim that Mythos has already found severe vulnerabilities in "every major operating system and web browser", including one that had gone undetected for nearly three decades.

How fast Mythos' capabilities have emerged is, as the name suggests, almost mythical in terms of scale. Just last month, Anthropic had announced that its previous generation Opus 4.6 model "is currently far better at identifying and fixing vulnerabilities than at exploiting them".

For example, Opus 4.6 turned the vulnerabilities it had found in Mozilla's Firefox 147 JavaScript engine (a programme that executes JavaScript code by taking human-readable JavaScript and converting it into machine code that the computer can run) into JavaScript shell exploits (or a successful attack that gives the user control over the command line, or shell, of the victim's environment through the browser's engine) only two times out of several hundred attempts.

Anthropic re-ran the same experiment as a benchmark for Mythos Preview. It developed working exploits 181 times, and achieved register control (taking control of the general behaviour of a CPU, or the brain of a computer system) 29 times, according to Anthropic. Non-experts, the company said, can also use Mythos Preview to find and exploit sophisticated vulnerabilities.

Engineers at Anthropic with no formal security training have asked Mythos Preview to find remote code execution vulnerabilities overnight, and woken up the following morning to a complete, working exploit.

"We did not explicitly train Mythos Preview to have these capabilities. Rather they emerged as a downstream consequence of general improvements in code, reasoning, and autonomy. The same improvements

that make the model substantially more effective at patching vulnerabilities also make it substantially more effective at exploiting them," the startup said in its April 7 note. All of these features have effectively prompted Anthropic to hit pause on its release.

Real-world testing

Earlier this month, the UK AI Security Institute (AISi), in its evaluation of Mythos Preview, flagged a marked jump in cyber capability, especially in structured testing environments. One of its headline findings is that the model was able to solve 73% of expert-level cybersecurity challenges in benchmark settings, far higher than earlier frontier models. These tests were designed to approximate real-world vulnerabilities and required a mix of technical depth and problem-solving ability.

Where earlier models often showed patchy results at higher difficulty levels, Mythos demonstrated a stronger ability to sustain performance across complex challenges. This suggests improvements in reasoning and planning, enabling the model to navigate layered cyber problems rather than just isolated exploits.

Crucially, the institute highlighted Mythos's growing "agentic" behaviour. It was able to execute long, multi-stage attack chains, indicating a shift from tool-like assistance to more autonomous capability. This raises concerns about how such systems could lower the barrier for less-skilled actors to conduct sophisticated cyber operations.

In one of the tests designed by the UK AISi, a 32-step corporate attack simulation spanning reconnaissance to full network takeover, Mythos became the first model to solve the task from start to finish in three out of its 10 attempts. AISi said it would take humans 20 hours to complete the same tasks.

Mythos breach

Even as it held back commercial deployment citing all these concerns, Anthropic parallelly announced Project Glasswing, which aims to assist companies that hope to use Mythos to step up cyber defences before it is widely released. Major software developers — including Apple, Nvidia, the Linux Foundation and CrowdStrike, as well as competitor Google were in this list.

What has now set the cat among the pigeons is a report by Bloomberg that the Mythos model was accessed by "a handful of users" in a private Discord chat on the day it was announced publicly, despite the restricted release. US Federal Reserve Chair Jerome Powell and Treasury Secretary Scott Bessent met with top American bank CEOs in a closed-door meeting earlier this month to discuss the cybersecurity risks posed by Mythos.

In the long run, Anthropic's researchers said they expect that defence capabilities will dominate: that the world will emerge more secure, with software better hardened — in large part by code written by these models. But the transitional period will be fraught. The leak of the software on the Discord chat and the question marks that it raises is a case in point.

Globally, most experts see the compelling need for an overarching consensus when it comes to governing AI. Even though regulatory approaches to the technology could differ from one jurisdiction to another, the need for global coordination on AI and an agreement on a set of standards and common guardrails is being articulated. Without that, this technology is going to be very difficult to contain and tough to enforce.

AI rethink in India?

Finance Minister Sitharaman has told banks to exercise a "high-degree" of vigilance and develop a coordination mechanism to respond to threats emerging from capabilities of this model.

Sofar, India has maintained that it wants to maintain a light touch regulatory framework for AI, looking to incentivise the potential economic aspects of the technology. The Mythos episode could force a rethink.

In the aftermath of Mythos, the IT Ministry set up the AI Governance and Economic Group — an inter-ministerial body — as the apex mechanism for coordinating AI policy and governance framework. A separate Technology and Policy Expert Committee was also created to provide expertise to the inter-ministerial group.

Common guardrails

Globally, most experts see the need for an overarching consensus when it comes to governing AI

Without that, this technology could become difficult to regulate