Project Glasswing: India’s Cybersecurity Agencies to Get Access to Anthropic’s Powerful AI

June 7, 2026 • vajiramandravi.com

Project Glasswing Latest News

- US AI company Anthropic is extending access to its restricted cybersecurity programme, Project Glasswing, to select organisations in India — including key government agencies responsible for protecting the country’s critical infrastructure.
- This follows Anthropic’s announcement that it would expand the programme from its initial US and UK participants to over 15 countries, with India being a significant addition.

What Is Project Glasswing and What Is Mythos

- **Anthropic** is one of the world’s leading AI companies, known for its AI model Claude.
- But it has also developed a far more powerful and restricted model called **Claude Mythos** — described as a **frontier AI model** capable of identifying critical software vulnerabilities at a level that could “fundamentally alter the balance between cyber attackers and defenders.”

- This is a significant claim. Most cybersecurity today depends on human experts finding software weaknesses before attackers do.
- A model that can do this at scale and speed could be a game-changer — but in the wrong hands, it could also be catastrophically dangerous. That is why Anthropic has kept Mythos strictly restricted and not publicly released.
- **Project Glasswing** is the controlled programme through which Anthropic shares access to **Mythos Preview** (the testing version) with a carefully vetted set of trusted organisations.
- Each organisation must meet Anthropic's security requirements before gaining access.

Which Indian Agencies Are Getting Access

- The following Indian government bodies are understood to have received — or are in line to receive — access to Mythos:
    - **I4C** — Indian Cyber Crime Coordination Centre
    - **CERT-In** — Indian Computer Emergency Response Team (the nodal agency for cybersecurity incidents in India)
    - **NCIIPC** — National Critical Information Infrastructure Protection Centre (falls under the National Security Advisor in the Prime Minister's Office)
    - **DIP** — Department of Telecommunications' Digital Intelligence Platform
- Additionally, some cybersecurity-focused research institutions have received access, and discussions are underway to extend it to cybersecurity and AI teams within India's largest IT services companies.
- The purpose is specific: NCIIPC and CERT-In requested access to use Mythos to identify vulnerabilities within India's banking and power infrastructure — before attackers can find and exploit them.

Why India Is Particularly Concerned

- India's concern about Mythos is two-sided — both as an opportunity and as a threat.

The Defensive Opportunity

- India's critical infrastructure — banking systems, power grids, telecom networks — is a high-value target for cyberattacks.
- A tool that can proactively find and fix software vulnerabilities in these systems before adversaries exploit them would be enormously valuable.

The Offensive Threat

- At the same time, India's government was also worried about what Mythos could do to India's systems in the wrong hands.
- Finance Minister Nirmala Sitharaman, in April 2026, held a high-level meeting with IT Minister Ashwini Vaishnaw to assess risks posed by Mythos to India's banking sector.
- The meeting resulted in concrete directions: the Indian Banks' Association (IBA) was asked to develop a coordinated institutional response mechanism, and banks were directed to engage top cybersecurity

professionals to continuously strengthen their defensive and monitoring capabilities.

The Geopolitical Dimension: Anthropic vs. the US Pentagon

- As per various reports, Anthropic is helping the **US National Security Agency (NSA)** deploy Mythos for **offensive cyber operations** — specifically to infiltrate networks of countries like China and Iran.
- This creates an awkward contradiction. Anthropic is simultaneously:
    - Providing Mythos to trusted allies (including India) for defensive cybersecurity.
    - Reportedly assisting the NSA with offensive cyber operations.
    - Fighting a legal battle with the US Department of Defense (which includes the NSA) over the boundaries of AI use.
        - The legal dispute arose because Anthropic drew a firm line — it refused to allow its Claude AI models to be used for mass surveillance of US citizens or lethal autonomous drones.
        - The Pentagon responded by labelling Anthropic a “supply-chain risk” — an unprecedented designation for a US company.
        - Anthropic has sued over this label.

Conclusion

- AI is becoming a national security asset. Governments are no longer just regulating AI — they are actively seeking to deploy frontier AI models for defence and cybersecurity. The race to access the most powerful AI tools is becoming part of geopolitical competition.
- Critical infrastructure protection is increasingly about software. Banking systems, power grids, and communications networks all run on software. Identifying and fixing vulnerabilities in that software is now as important as physical security.
- AI companies are navigating unprecedented ethical and geopolitical pressure. Anthropic’s tension with the Pentagon illustrates that even safety-focused AI companies face difficult choices about how their models are used — and by whom.

Source: IE
