

India Locked Out

Experts warn Western AI benchmarks leave India's public infra exposed, writes Tanya Pandey

The most powerful artificial intelligence models are increasingly being tested behind closed doors before they reach the wider world. India now wants a way into that first line of testing.

Policy experts argue that India should create a formal mechanism giving government agencies, technical institutions, and independent researchers early access to frontier AI models in controlled environments for safety and red-team testing. Findings could be shared confidentially with developers so vulnerabilities can be addressed before extensive deployment.

The demand comes against a widening gap. India has limited homegrown frontier AI model capabilities and no formal system to secure early access to the most capable models being developed overseas for independent safety and cybersecurity testing, leaving it with limited ability to assess them against risks specific to the country.

"If India is consistently brought into the conversation only after the first round of frontier-model testing, we risk becoming a rule-taker rather than a rule-shaper," said Kazim Rizvi, founder, The Dialogue, a policy think tank.

The first round of evaluations can influence which risks are prioritised, what is considered acceptable model behaviour and the safeguards that eventually become industry standards, Rizvi said. Coming in later could also leave Indian institutions dependent on safety assessments conducted elsewhere for systems that may eventually be deployed across the country.

This is of particular importance because India's technology environment is different from those in which much of the early testing takes place.

"India has a highly diverse language landscape, rapidly digitising public services and distinctive cybersecurity risks," Rizvi said.

A model that performs safely in English under Western testing conditions could behave differently when deployed across Indian languages or connected to financial, identity and public digital systems, he said. "The bigger issue is lack of access, not just coming in late," said Nikhil Narendran, partner, Trilegal.

India needs to test frontier models against risks specific to its context, particularly those involving critical infrastructure. "Without early access, we have less room to put safeguards in place, while other states or state-sponsored

actors could exploit that window against our interests," Narendran said.

The country should establish a formal framework for early access and independent testing. The challenge, he added, would be getting frontier AI companies to participate, given India's limited leverage. The proposed mechanism would not necessarily mean unrestricted access to powerful models. Rizvi suggested a trusted evaluation programme in which designated government agencies, technical institutions, and independent experts could test models in controlled environments under strict confidentiality, information-security, and conflict-of-interest safeguards.

Testing could focus on risks particularly relevant to India, including attacks on critical infrastructure, automated cyber operations, misinformation, fraud and vulnerabilities created when AI systems are integrated into public digital infrastructure.

This could also give developers an incentive to participate. Serious vulnerabilities identified during testing could first be shared confidentially with the company so they can be addressed before wider deployment. Regulators could receive appropriately classified assessments without requiring sensitive technical details to be made public, Rizvi said.

The objective, experts said, would be to ensure that India can independently understand the systems being deployed across its economy rather than relying entirely on evaluations conducted overseas.

There is a difference between not being publicly identified as an early tester and being completely excluded from testing. "The limited public visibility of Indian organisations among initial testers does not by itself demonstrate that India has been excluded," said Apeksha Kaushik, senior principal analyst, Gartner. Frontier-model testing is often conducted under confidentiality agreements, she said.

India should nevertheless create more formal pathways for its organisations to participate in independent red-team programmes, cyber ranges, responsible-disclosure initiatives and public-private safety partnerships, Kaushik said.

For India, the issue is becoming more consequential as the global debate turns towards whether frontier AI development should become more gradual. A slowdown would affect countries differently, depending on how far they have progressed in building advanced models, computing infrastructure, and research capacity.

"India cannot afford to interpret calls for caution as a reason to step back from frontier AI development," Rizvi said. An uneven slowdown could preserve the advantage of countries that already possess advanced models, compute infrastructure, capital and research talent, he said.

As most frontier AI researchers warn of doomsday scenarios, the battle to control superintelligence has become our modern existential crisis of almost Shakespearean proportions

The AI Genie's Out: From Annoyance to Apocalypse



IPO PLANS



OpenAI: Moved IPO plans to 2027 or later



Anthropic: Confidential S1 filing made on June 1, 2026; potential IPO preparation underway