

Why are Indian firms racing to build local AI? | Explained

What are the objectives of AI Kosha? How will the Common Compute facility providing Graphics Processing Units help startups, researchers and developers? How many GPUs have been commissioned? What is the roadmap? Why are translation models important?

Published - March 09, 2025 06:01 am IST

AROON DEEP



Artificial Intelligence Machine Learning Large Language Model Technology | Photo Credit: Getty Images

The story so far: On March 6, the Ministry of Electronics and Information Technology's **IndiaAI Mission** launched **AI Kosha**, a national dataset platform that will gather non-personal data from various ministries and departments in order to develop Artificial Intelligence models and tools. It also launched the **Common Compute portal** to allow startups and academia to access a shared pool of Graphics Processing Units (GPUs) to build and run AI models and tools.

What is the IndiaAI Mission?

The **IndiaAI portal** was initially set up as a joint initiative between the Ministry of Electronics and Information Technology and the Indian tech industry association Nasscom in 2023. The portal has evolved into a mission led entirely by the Union government, announced last March with an outlay of ₹10,371.92 crore. The mission was approved by the Union Cabinet amid the rapid proliferation and development of generative AI models like ChatGPT and an increased global enthusiasm to develop AI. The key objectives of the mission were: "Making AI in India and Making AI Work for India".

While the mission works on many interventions and resources for private companies and academia to research and build AI, a key one that has emerged in recent weeks has been the Common Compute facility to provide GPUs to startups that may not be able to buy them independently in a large volume.

GPUs, also known as graphics cards, are a key component in many PCs and phones that allow these devices to render images and graphics. However, AI researchers have long used GPUs to run enormously complex calculations involving massive volumes of data. Generative AI models, which can only be created (“trained”) and run when they have access to such computing resources, have become a major use case of GPUs. This is what has enabled firms like Nvidia to foray far beyond video gaming to become an essential component of AI infrastructure around the world. However, GPUs are expensive. Blackwell, a GPU made by Nvidia, is reportedly sold out until this October. As such, the IndiaAI Mission has started empanelling data centres in India that have GPUs to provide access to their computing resources to startups and researchers. This access will be handled by the mission, which will decide which entities will get how much computing resource access, and for how long. As of now, 14,000 GPUs have been acquired and commissioned for use by empanelled data centres so far, according to IT Minister Ashwini Vaishnaw, with more to be acquired on a quarterly basis.

What is AI Kosha?

AI Kosha is the implementation of another pillar of the mission: the IndiaAI Datasets Platform. One of the main issues with building AI models from scratch — or to some extent even adapting existing models to India — is that they are largely trained on Western models that are heavily biased towards English and developed countries. The fact that many internet users are navigating the web and contributing to discussions online in English even if it is not their main language in India — as The Hindu has reported — has been a major challenge in putting together such data. The AI Kosha is the beginning of an attempt, therefore, to make India-specific data much more widely available. The datasets that have so far been uploaded — such as translation models between several “pairs” of Indian languages — are an effort to diminish this bias by giving researchers and firms building AI tools in India wider access to local content.

Many open source AI models would theoretically allow an Indian firm to build their own powerful generative AI tools, such as chatbots. But local data, policymakers and developers have argued, are critical to ensure that such deployments are actually useful over what foreign firms like OpenAI and Google already offer.

What else does the IndiaAI Mission work on?

The mission has seven pillars overall. While the Common Compute facility and AI Kosha have been under particular focus due to developments like DeepSeek — the Chinese AI model that was trained at a far lower cost than competing models from the U.S. — there are other aspects that policymakers have targeted.

One is the AI Safety Institute of India, which is yet to be formally launched. Such institutes are being created or accredited around the world to develop guardrails around unintended consequences of AI deployments, and to study the risks of the mushrooming use of AI across the industry and government. These include everything from making sure that a chatbot doesn't give away a recipe for a bomb (for example), to making sure that specialised AI tools, like those used in biomedical research, are reliable and can "explain" their output.

Other pillars include the IndiaAI Innovation Centre, which aims to create "domain-specific foundation models"; an application development initiative to help develop use-cases that can be commercialised or put to productive use; Futureskills, an initiative to develop AI labs in smaller cities; and startup financing.

A key effort that dovetails with many of these pillars is the request for proposals put out by the mission to develop an indigenously developed foundation model. Building a foundational model — the main program that a generative AI tool runs on — is highly resource-intensive and accounts for the bulk of developing AI models. The Union government is keen on facilitating this process in a matter of months, Mr. Vaishnav has said, and 67 applications have already been received by the mission, which is evaluating these at present.

Published - March 09, 2025 06:01 am IST