

ADVERTISEMENT

News / Explained / Explained Sci-Tech / Explained: GPT-4o, OpenAI's newest AI model that makes ChatGPT smarter and free for all

Explained: GPT-4o, OpenAI's newest AI model that makes ChatGPT smarter and free for all

GPT-4o will be a big boost for free users, as they will now get GPT-4-level intelligence on a much faster and efficient ChatGPT.

Written by [Bijin Jose](#)

Bengaluru | Updated: May 14, 2024 18:53 IST

 NewsGuard

EXPRESS EXPI

What are Ind
Chabhar po

GPT-4o will be made available to the public in stages. Text and image capabilities are already rolling out on ChatGPT, with some services available to free users. (Via OpenAI)

OpenAI introduced its latest large language model (LLM) called GPT-4o on Monday (May 13), billing it as their fastest and most powerful AI model so far. The company claims that the new model will make ChatGPT smarter and easier to use.

Until now, OpenAI's most advanced LLM was the GPT-4, which was only available to paid users. However, the GPT-4o will be freely available.

What is GPT-4o?

GPT-4o ("o" stands for "Omni" here) is being seen as a revolutionary AI model, which has been developed to enhance human-computer interactions. It lets users input any combination of text, audio, and image and receive responses in the same formats. This makes GPT-4o a multimodal AI model – a significant leap from previous models.

Describing the new model, OpenAI CTO Mira Murati said that it was the first time

that OpenAI was making a huge step forward when it came to ease of use.

Based on the live demos, GPT-4o seems like ChatGPT transformed into a digital personal assistant that can assist users with a variety of tasks. From real-time translations to reading a user's face and having real-time spoken conversations, the new model is far ahead of its peers.

EXPRESS EXP

What are Ind
Chabhar po

GPT-4o is capable of interacting using text and vision, meaning it can view screenshots, photos, documents, or charts uploaded by users and have conversations about them. OpenAI said that the new updated version of ChatGPT will also have updated memory capabilities and will learn from previous conversations with users.

What is the technology behind GPT-4o?

LLMs are the **backbone of AI chatbots**. Large amounts of data are **fed** into these models to make them capable of learning things themselves.

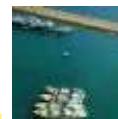
ADVERTISEMENT

Read | Google introduces Gemma open source AI models: What does it mean for responsible AI?

Unlike its predecessors that required multiple models to handle different tasks, GPT-4o uses a single model trained end-to-end across various modalities – text, vision, and audio. To illustrate this, Murati explained the voice mode on previous models, which was a combination of three different models – transcription, intelligence, and text-to-speech. With GPT-4o, all of this happens natively.

Essentially, this means the GPT-4o comes with an integration that allows it to process and understand inputs more holistically. For example, GPT-4o can understand tone, background noises, and emotional context in audio inputs once. These abilities were a big challenge for earlier models.

EXPRESS EXPI

What are Ind
Chabaha po

When it comes to features and abilities, GPT-4o excels in areas like speed and efficiency as it responds to queries as fast as a human does in conversation, in around 232 to 320 milliseconds. This is a big leap over previous models, which came with response times of up to several seconds.

It comes with multilingual support, and shows significant improvements in handling non-English text, making it more accessible to a global audience.

Ailing Ex-Bengal CM Buddhadeb Bhattacharjee Comes To Life With AI...



The GPT-4o also features enhanced audio and vision understanding. During the demo session at the live event, ChatGPT solved a linear equation in real-time when the user was writing it on paper. It could gauge the emotions of the speaker on camera and identify objects.

Why does it matter?

It comes at a time when the AI race is intensifying, with tech giants Meta and Google working towards building more powerful LLMs and bringing them to

various products. GPT-4o could be beneficial for [Microsoft](#), which has invested billions into OpenAI, as it can now embed the model in its existing services.

The new model also came a day ahead of the [Google I/O developer conference](#) where Google is expected to announce new updates to its [Gemini](#) AI model. To GPT-4o, Google's Gemini is also expected to be multimodal. Further, at the Worldwide Developers Conference in June, announcements on incorporating AI in iPhones or [iOS](#) updates are expected.

EXPRESS EXPI

What are Ind
Chabhar po

When will GPT-4o be available?

GPT-4o will be made available to the public in stages. Text and image capabilities are already rolling out on ChatGPT, with some services available to free users. Audio and video functionalities will come gradually to developers and selected partners, ensuring that each modality (voice, text-to-speech, vision) meets the necessary safety standards before full release.

What are GPT-4o's limitations and safety concerns?

Although it is claimed to be the most advanced model, GPT-4o is not without limitations. On its official blog, OpenAI said that GPT-4o is still in the early stages of exploring the potential of unified multimodal interaction, meaning certain features like audio outputs are initially accessible in a limited form only, with preset voices.

The company said that further development and updates are necessary to fully realise its potential in handling complex multimodal tasks seamlessly.

When it comes to safety, OpenAI said that GPT-4o comes with built-in safety measures, including "filtered training data, and refined model behaviour post training". The company claimed that the new model has undergone extensive safety evaluations and external reviews, focussing on risks like cybersecurity, misinformation, and bias.

As of now, while GPT-4o scores only a Medium-level risk across these areas, OpenAI said that continuous efforts are in place to identify and mitigate emerging risks.