ChatGPT and Whisper APIs debut, allowing devs to integrate them into apps

reader comments

26 with

Share this story

On Wednesday, OpenAI announced the availability of developer APIs for its popular ChatGPT and Whisper AI models that will let developers integrate them into their apps. An API (application programming interface) is a set of protocols that allows different computer programs to communicate with each other. In this case, app developers can extend their apps' abilities with OpenAI technology for an ongoing fee based on usage.

Introduced in late November, ChatGPT generates coherent text in many styles. Whisper, a speech-to-text model that launched in September, can transcribe spoken audio into text.

In particular, demand for a ChatGPT API has been huge, which led to the creation of an unauthorized API late last year that violated OpenAI's terms of service. Now, OpenAI has introduced its own API offering to meet the demand. Compute for the APIs will happen off-device and in the cloud.

OpenAI calls its new ChatGPT API model "gpt-3.5-turbo," which supersedes its previous "best" LLM API, "text-davinci-003." It is priced at $0.002 per 1,000 tokens (about 750 words), which

OpenAI says is about 10 times cheaper than its existing GPT-3.5 models. “Through a series of system-wide optimizations, we’ve achieved 90% cost reduction for ChatGPT since December,” writes OpenAI on its API announcement page.

Advertisement

Enlarge / An example of a call to the ChatGPT API.
Early users of the ChatGPT API include Snapchat, with its “My AI” bot we covered Monday; an educational platform called Quizlet that reportedly helps students study; and Instacart, which plans to add “Ask Instacart” later this year to let customers ask about food.

The Whisper API, which is based on the open source whisper-large-v2 model, is available at a price of $0.006 per minute. It accepts inputs in M4A, MP3, MP4, MPEG, MPGA, WAV, and WEBM formats and can translate audio to text or transcribe at a rate comparable to a skilled human transcriptionist, even with difficult audio.

Responding to privacy concerns about collecting data from customers, OpenAI says it has modified its terms of service (TOS) to no longer use data submitted for service improvements, including future model training. Also, it has a default 30-day data-retention policy for API users, and the TOS emphasizes that users own the input and output of the AI models.

For its AI models, OpenAI is also now offering dedicated cloud instances running on Azure “for users who want deeper control over the specific model version and system performance.”

The ChatGPT and Whisper APIs are available starting today. Usage requires an OpenAI account and a payment method.

https://arstechnica.com/?p=1920999