

Hello GPT-4o

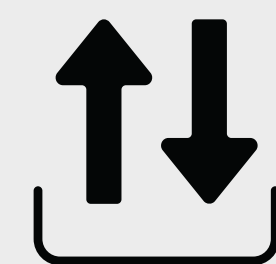


The new model that can reason across audio, vision, and text in real time announced by OpenAI on May 13, 2024.

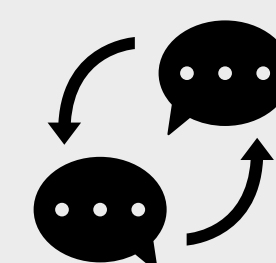
MORE ADVANCED



Coding



Multimodal Inputs/Outputs



Real-time Interaction

MORE ACCESSIBLE

FREE

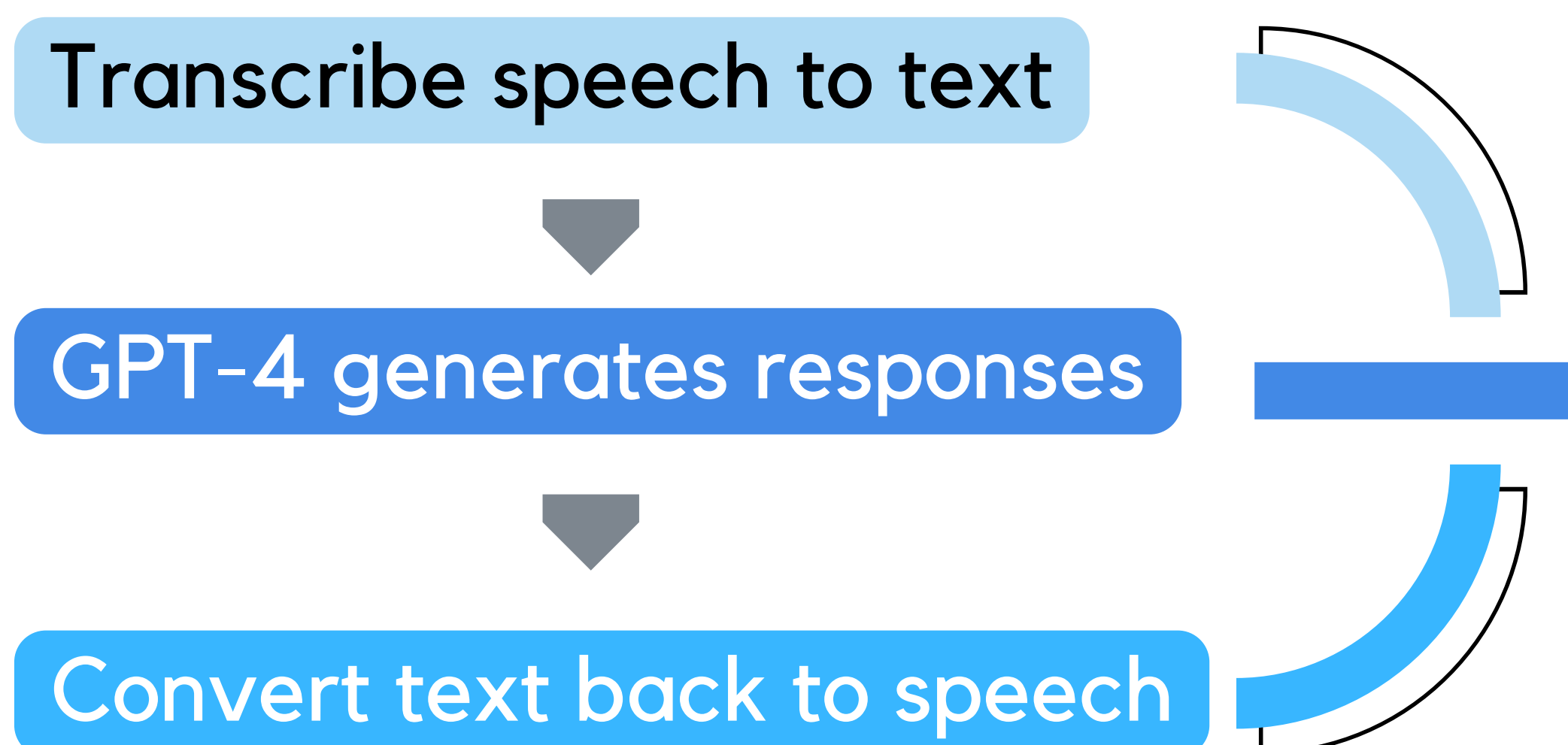
GPT-4o is free for OpenAI users.

MORE FLEXIBLE

GPT-4o accepts as input any combination of text, audio, image, and video and generates any combination of text, audio, and image outputs.

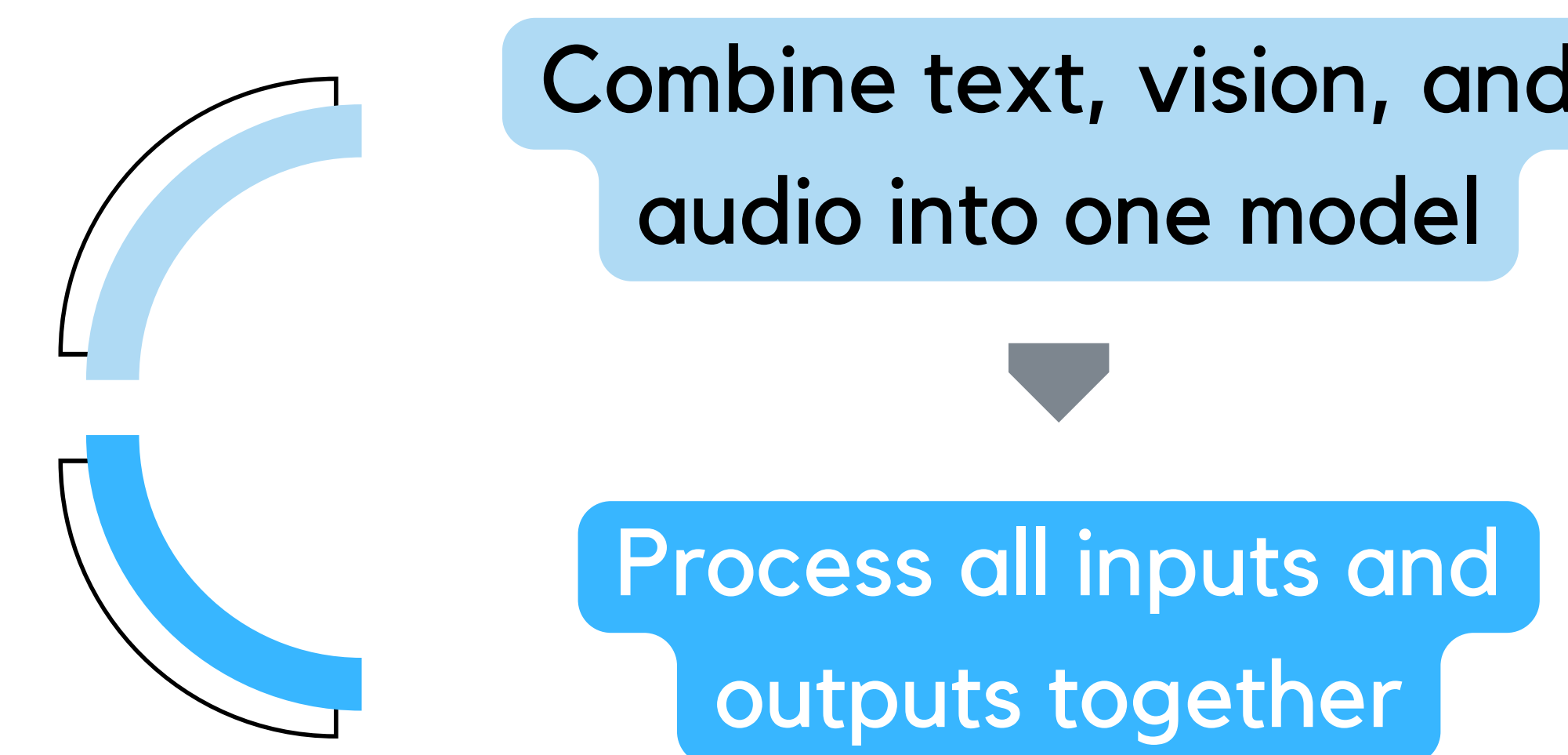
One Big Step Forward

in Voice Interactions



GPT-4

GPT-4o



Cons: GPT-4 misses out on emotional inflections of the speaker, and other details

Pros: This unified approach allows it to better understand and generate nuanced responses.

Use Cases for GPT-4o

By OpenAI

- **Customer Service:** GPT-4o can handle customer service inquiries, including making phone calls to resolve issues.
- **Creative Content:** It can generate movie posters, caricatures, and custom fonts based on user inputs.
- **Meeting Transcriptions and Summaries:** The model can transcribe and summarize meetings with multiple speakers.
- **Real-time Translation:** It can translate conversations in real-time.
- **Enhanced Interaction:** GPT-4o provides more natural, human-like interactions, including recognizing tone and emotion.

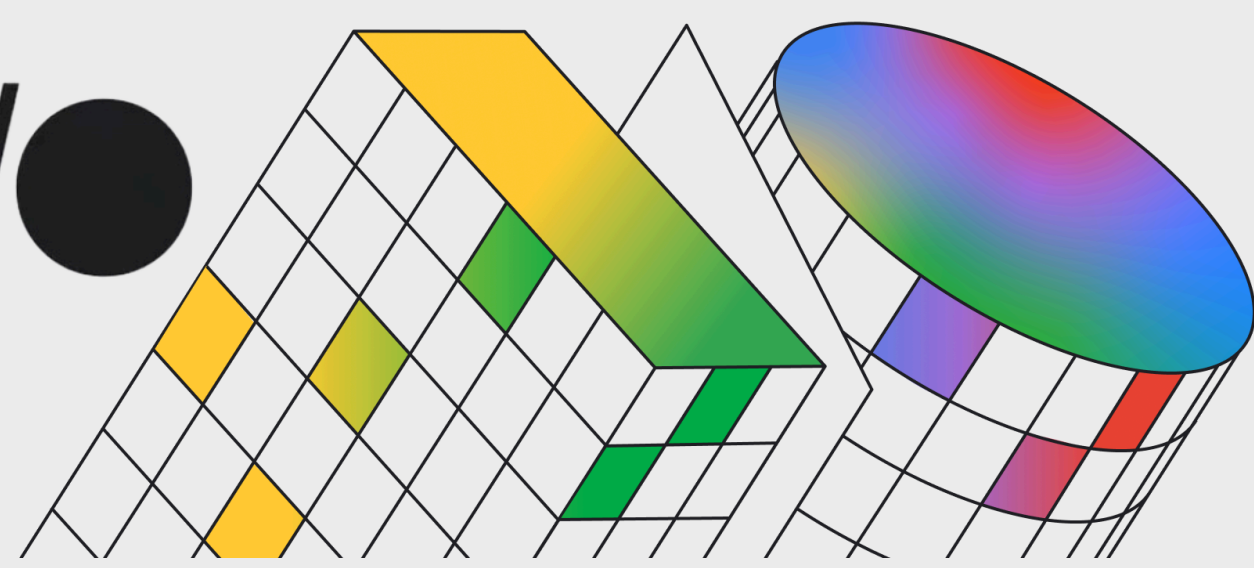


**GPT-4o
Announcement**

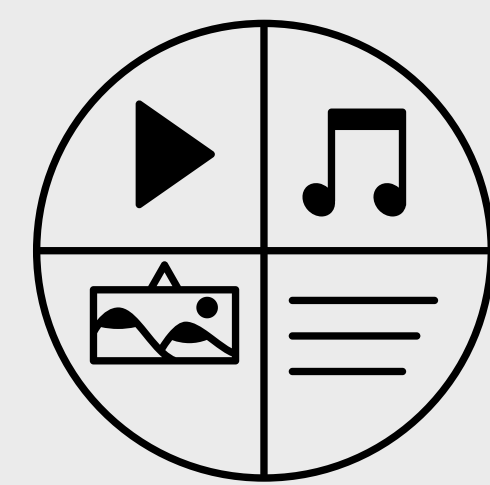


Google I/O

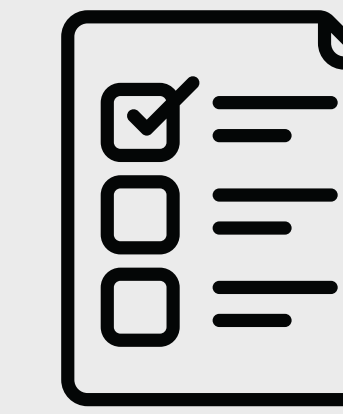
May 14, 2024



Gemini 1.5

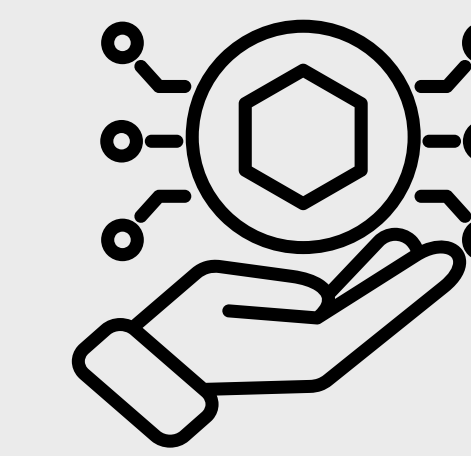


Allow multiple types of input, including text, images, audio, and video.



Tasks It Can Perform:

- transcribe and analyze videos;
- understand and process audio;
- reason across different types of data simultaneously.



Support a context window of up to 1 million tokens.

Compared to 128K context window by GPT-4o.



Prompts for Education By Google



Generative AI Course for Educators

The latest announcement from Google about the AI launches on May 14, 2024.

Google NotebookLM

Google NotebookLM now is powered by Gemini 1.5.

- It allows users to upload a number of different documents;
- It acts as a virtual research assistant, capable of summarizing facts, explaining complex ideas, and brainstorming new connections based on the provided sources.
- Users can ground the AI in specific documents to limit hallucinations.

Ask Photos

Gemini is integrated in Google Photos, called 'Ask Photos'.

This feature allows users to search for specific images in their gallery using natural language queries.

The AI processes queries in the cloud, ensuring privacy and security, and does not store queries or use personal data for training other AI models.

Google For You

AI-powered search results is introduced, called the **Search Generative Experience**. It aims to provide direct answers within search results, reducing the need for users to visit multiple web pages.

The future of web interactions may see Google increasingly centralizing information delivery, raising concerns among content creators and web publishers.

Gemini in Gmail

Gmail is introducing AI-powered email summaries to help users manage their inboxes more efficiently. This new feature leverages generative AI to provide concise summaries of email threads. It's part of a broader initiative by Google to integrate AI into its products.

Gemini Nano

Google announced the integration of its Gemini Nano AI model into the Chrome desktop client, starting with version 126. This model, optimized for performance on various hardware, will enable developers to incorporate AI features like translation, captioning, and transcription directly into the browser.

Project Astra

Google shared the progress of developing universal AI agents with Project Astra (advanced seeing and talking responsive agent).

It is a real-time, multimodal AI assistant that can see the world, knows what things are and where you left them, and can answer questions or help you do almost anything.

Gen Media Models

Google introduced Veo, a model for generating high-definition video, and Imagen 3, a text-to-image model. They also shared a new demo recordings created with the Music AI Sandbox, Lyria.

Veo might not be as impressive as OpenAI Sora, but possibly will be released to the public earlier than Sora.



Watch Google I/O '24