

You Can Now Talk to ChatGPT With OpenAI's Real-Time Voice and Video Capabilities



Mark your calendar for Mediaweek, October 29-30 in New York City. We'll unpack the biggest shifts shaping the future of media—from tv to retail media to tech—and how marketers can prep to stay ahead. [Register](#) with early-bird rates before sale ends!

OpenAI announced its new flagship model GPT-4o (letter “o” for “Omni”), boasting enhanced conversational abilities. This upgrade provides all ChatGPT users with faster, real-time voice and video interaction capabilities.

The new chatbot, featuring multimodal capabilities spanning visual, audio and text interactions, allows people to use a phone's camera to allow the AI assistant to read written text as well as detect a person's emotions, including the ability to interrupt the AI assistant. Additionally, users can interrupt the AI assistant during interactions.

"The special thing about GPT-4o is that it brings GPT-4 level intelligence to everyone, including our free users," OpenAI's CTO Mira Murati said during a livestream presentation.

During the livestream, one user asked ChatGPT for a real-time tutorial on how to take deep breaths, with the AI model picking up on the user's emotional state such as breathing too fast. In another demo, ChatGPT narrated a story in various emotive styles, including dramatic to robotic tones, and even singing. In a third demo, the user asked ChatGPT to look at a math equation and assist the person in solving it, instead of the AI chatbot providing an answer.

"This is the end of prompting and beginning of conversing with AI models," Elav Horwitz, evp of global head of applied innovation at McCann Worldgroup, told ADWEEK.

OpenAI's improved capabilities announcement comes a day before Google's I/O developer conference, shooting down reports about OpenAI's search product launch to rival Google and Perplexity. For marketers, OpenAI's latest version of ChatGPT enables a seamless integration of gen AI within creative processes such as brainstorming and generating ad copy.

"This technology has the potential to transform the chatbots used by brands on their apps and websites, making AI interactions much more human-like, and mimicking personal shoppers or customer support representatives," said William Chen, director of product management, AI and emerging tech at Agora. "By using the real-time engagement capability to understand and respond through multiple modalities ... AI models like GPT-4o can offer personalized and empathetic on-the-spot customer service or entertainment at scale for marketers."

However, this might deter certain brands that are already hesitant about adopting AI, said Cory Treffiletti, AI startup Rembrand's chief marketing officer.

"People can get the AI to say what they want," said

Treffeletti. “And brands want more control over their interaction with consumers.”

OpenAI’s new chatbot also operates seamlessly across languages, exemplified by its ability to translate between English and Italian.

The GPT-4o capabilities will be available in the coming weeks, with its text and image capabilities starting to roll out today in ChatGPT, per OpenAI’s blog post. The company also announced the release of a desktop version of ChatGPT, launching first for Mac users, with access available to paid users starting today.

Recommended articles

<https://www.adweek.com/media/you-can-now-talk-to-chatgpt-with-openais-real-time-voice-and-video-capabilities/>