Realtime API Beta

⊡ Copy page

Build low-latency, multi-modal experiences with the Realtime API.

The OpenAI Realtime API enables you to build low-latency, multi-modal conversational experiences with expressive voice-enabled models. These models support realtime text and audio inputs and outputs, voice activation detection, function calling, and much more.

The Realtime API uses GPT-4o and GPT-4o-mini models with additional capabilities to support realtime interactions. The most recent model snapshots for each can be referenced by:

`gpt-4o-realtime-preview-2024-12-17`

`gpt-4o-mini-realtime-preview-2024-12-17`

Dated model snapshot IDs and more information can be found on the models page here.

Get started with the Realtime API

Learn to connect to the Realtime API using either WebRTC (ideal for client-side applications) or WebSockets (great for server-to-server applications). Once you connect to the Realtime API, learn how to use client and server events to build your application.

🖥️ **Connect to the Realtime API using WebRTC**
To interact with Realtime models in web browsers or client-side applications, we recommend connecting via WebRTC. Learn how in this guide!

⚡ **Connect to the Realtime API using WebSockets**
In server-to-server applications, you can connect to the Realtime API over WebSocket as well. Learn how in this guide!

📖 **Realtime model capabilities**
Learn to use the Realtime API's evented interface to build applications using Realtime models. Learn to manage the Realtime session, add audio to model conversations, send text generation requests to the model, and make function calls to extend the capabilities of the model.

</> **Python SDK**
Beta support in the Python SDK for connecting to the Realtime API over a WebSocket.

</> **Full API reference**
Check out the API reference for all available interfaces.

Example applications

Check out one of the example applications below to see the Realtime API in action.

Realtime Console
To get started quickly, download and configure the Realtime console demo. See events flowing back and forth, and inspect their contents. Learn how to execute custom logic with function calling.

Client-side tool calling
Built with Cloudflare Workers, an example application showcasing client-side tool calling. Also check out the tutorial on YouTube.

Partner integrations

Check out these partner integrations, which use the Realtime API in frontend applications and telephony use cases.

LiveKit integration guide
How to use the Realtime API with LiveKit's WebRTC infrastructure.

Twilio integration guide
Build Realtime apps using Twilio's powerful voice APIs.

Agora integration quickstart
How to integrate Agora's real-time audio communication capabilities with the Realtime API.