

OpenAI Platform

< Models



GPT-5.2 pro

Default



Version of GPT-5.2 that produces smarter and more precise responses.

Compare

Try in Playground

Reasoning



Speed



Price

\$21 · \$168

Input



Output



GPT-5.2 pro is available in the Responses API only to enable support for multi-turn model interactions before responding to API requests, and other advanced API features in the future. Since GPT-5.2 pro is designed to tackle tough problems, some requests may take several minutes to finish. To avoid timeouts, try using background mode. GPT-5.2 pro supports reasoning.effort: medium, high, xhigh.

❖ 400,000 context window

➡ 128,000 max output tokens

📅 Aug 31, 2025 knowledge cutoff

💡 Reasoning token support

Pricing

Pricing is based on the number of tokens used, or other metrics based on the model type. For tool-specific models, like search and computer use, there's a fee per tool call. See details in the [pricing page](#).

Text tokens

Per 1M tokens • Batch API price 

Input

\$21.00

Output

\$168.00

Quick comparison

Input Output

GPT-5.2 pro	\$21.00
-------------	---------

o3-pro	\$20.00
--------	---------

GPT-5.2	\$1.75
---------	--------

Modalities

 **Text**
Input and output

 **Image**
Input only

 **Audio**
Not supported

 **Video**
Not supported

Endpoints

 **Chat Completions**
v1/chat/completions

 **Responses**
v1/responses

 **Realtime**
v1/realtime

 **Assistants**
v1/assistants

 **Batch**
v1/batch

 **Fine-tuning**
v1/fine-tuning

 **Embeddings**
v1/embeddings

 **Image generation**
v1/images/generations

 **Videos**
v1/videos

 **Image edit**
v1/images edits

 **Speech generation**
v1/audio/speech

 **Transcription**
v1/audio/transcriptions

 **Translation**
v1/audio/translations

 **Moderation**
v1/moderations

 **Completions (legacy)**
v1/completions

Features

 **Streaming**
Supported

 **Function calling**
Supported

Structured outputs

 Not supported

 **Fine-tuning**
Not supported

 **Distillation**
Not supported

Tools

Tools supported by this model when using the Responses API.

 **Web search**
Supported

 **File search**
Supported

 **Image generation**
Supported

 **Code interpreter**
Not supported

 **Computer use**
Not supported

 **MCP**
Supported

Snapshots

Snapshots let you lock in a specific version of the model so that performance and behavior remain consistent. Below is a list of all available snapshots and aliases for GPT-5.2 pro.

 **gpt-5.2-pro**
↳ gpt-5.2-pro-2025-12-11

gpt-5.2-pro-2025-12-11

Rate limits

Rate limits ensure fair and reliable access to the API by placing specific caps on requests or tokens used within a given time period. Your usage tier determines how high these limits are set

and automatically increases as you send more requests and spend more on the API.

TIER	RPM	TPM	BATCH QUEUE LIMIT
Free		Not supported	
Tier 1	500	30,000	90,000
Tier 2	5,000	450,000	1,350,000
Tier 3	5,000	800,000	50,000,000
Tier 4	10,000	2,000,000	200,000,000
Tier 5	10,000	30,000,000	5,000,000,000