Pricing and Setting Usage Limits for OpenAI

Making an AI Application with OpenAI API

Why set usage limits?

- Don’t want to be surprised by a big bill
- OpenAI API charges per use. Can grow quickly
- Pricing can grow quickly if your application goes viral

Current Pricing

As of 7/23/23

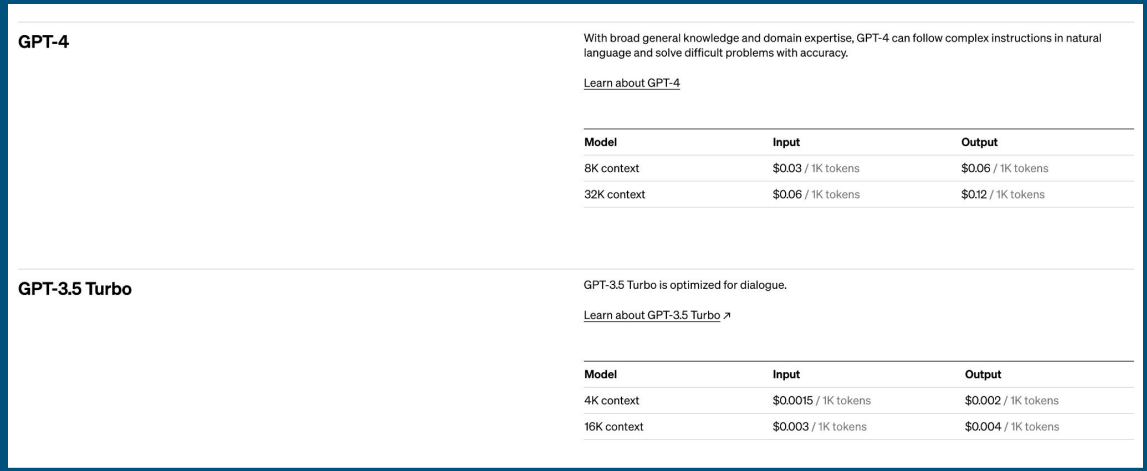
1 token = 1 word

For GPT 3.5 4k Context

For 10,000 words in and out, it costs $0.035

For 100,000 words in and out it costs for $0.35

For 1,000,000 words in and out it costs $3.5



GPT-4

With broad general knowledge and domain expertise, GPT-4 can follow complex instructions in natural language and solve difficult problems with accuracy.

Learn about GPT-4

Model	Input	Output
8K context	$0.03 / 1K tokens	$0.06 / 1K tokens
32K context	$0.06 / 1K tokens	$0.12 / 1K tokens

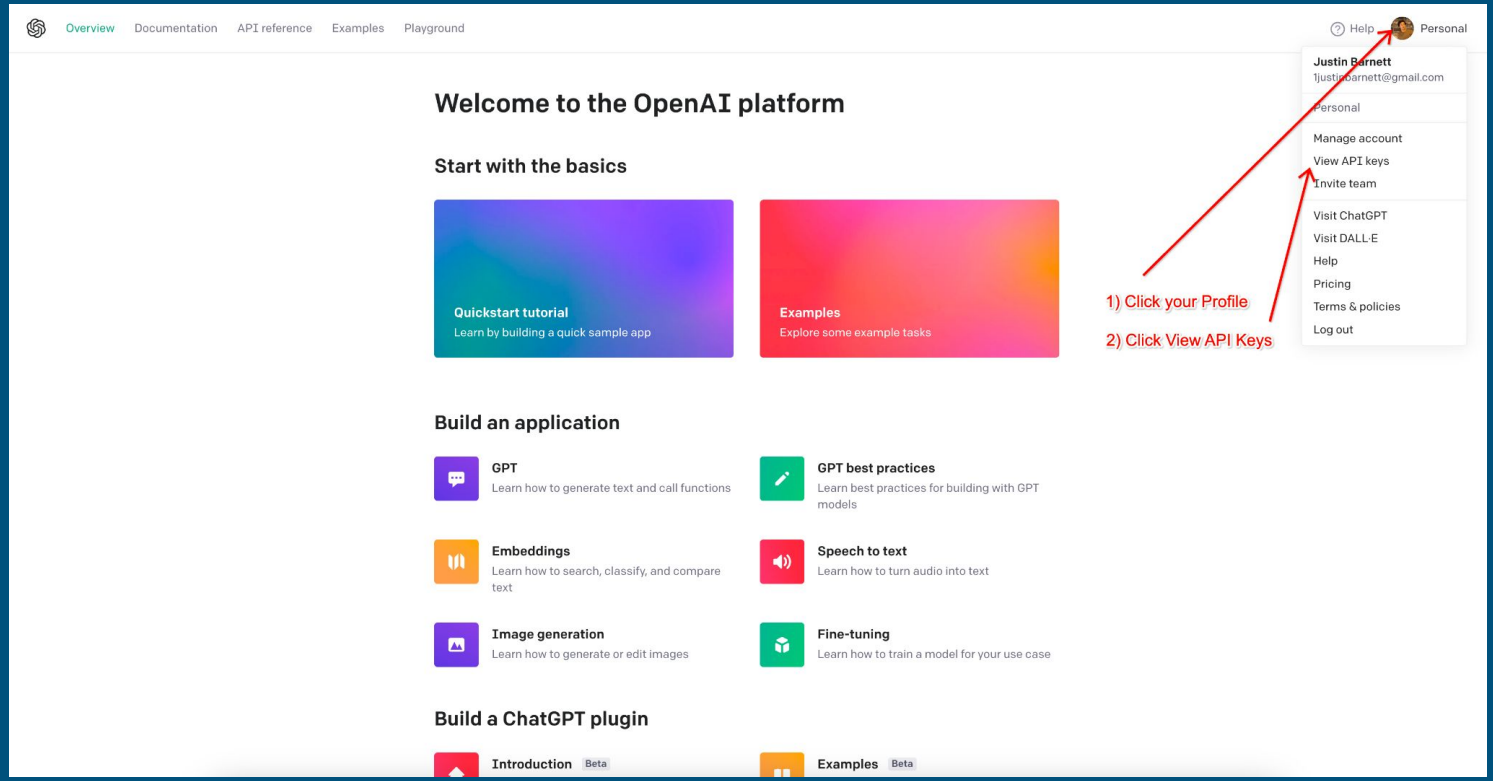
GPT-3.5 Turbo

GPT-3.5 Turbo is optimized for dialogue.

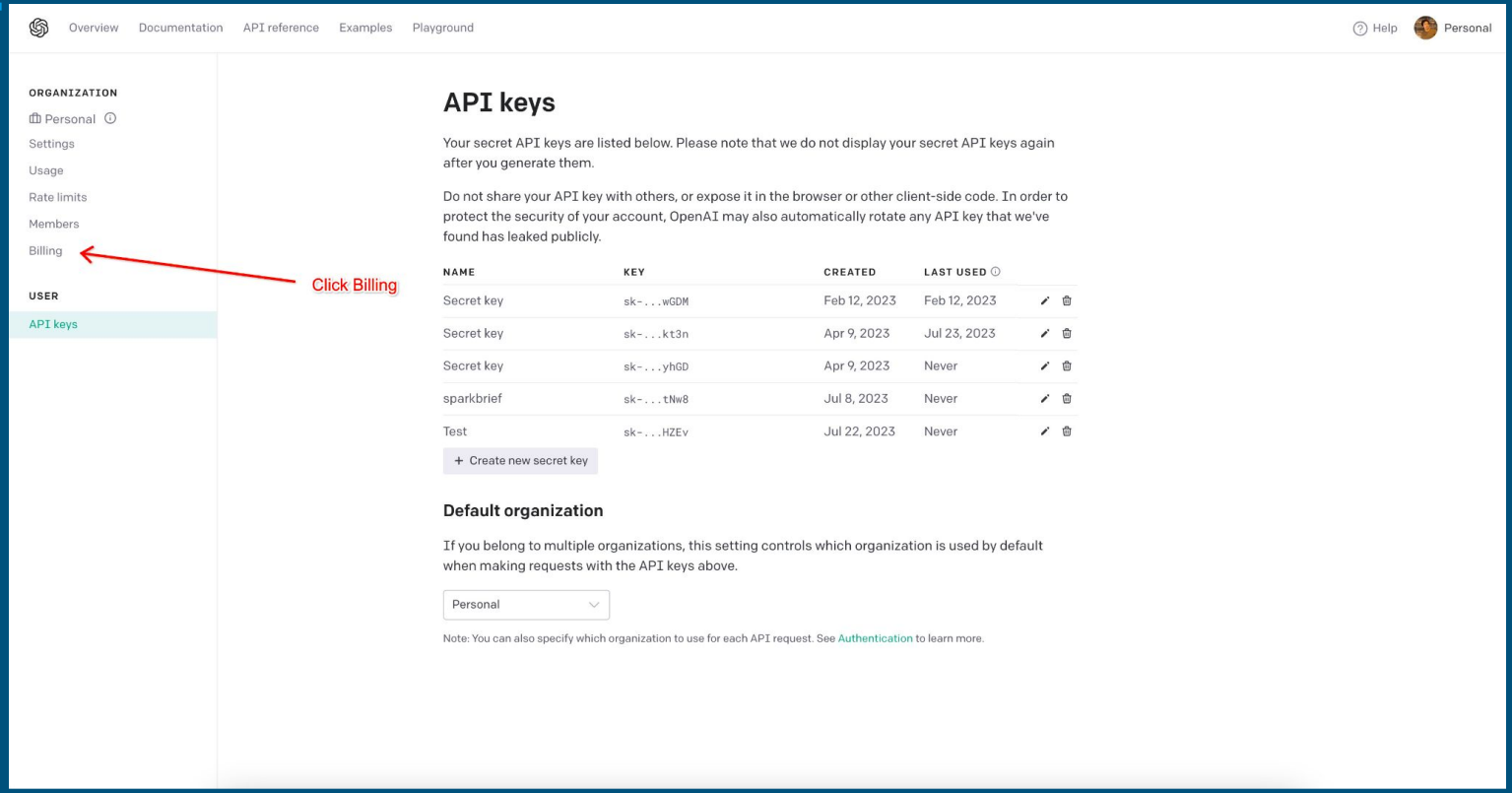
Learn about GPT-3.5 Turbo ↗

Model	Input	Output
4K context	$0.0015 / 1K tokens	$0.002 / 1K tokens
16K context	$0.003 / 1K tokens	$0.004 / 1K tokens

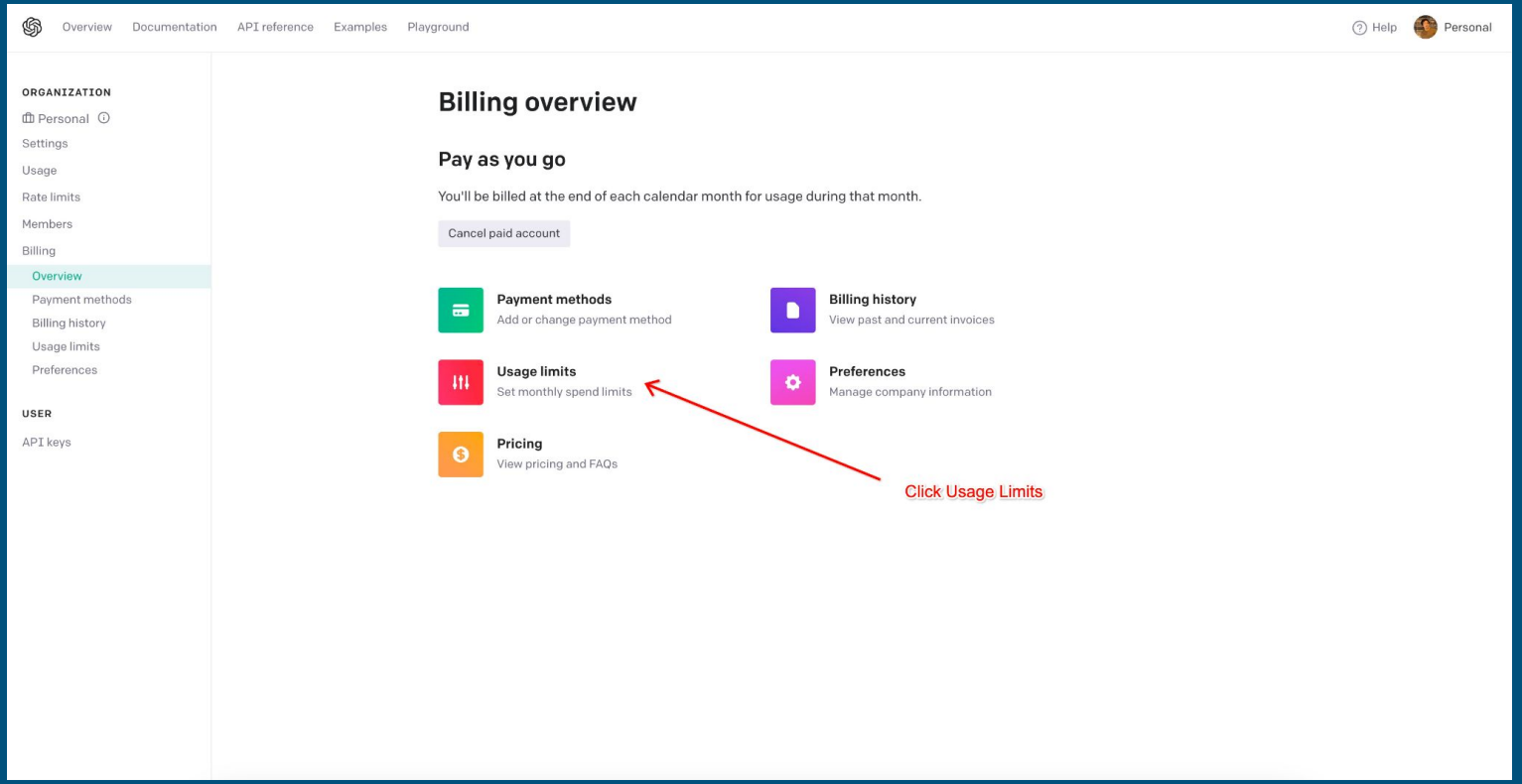
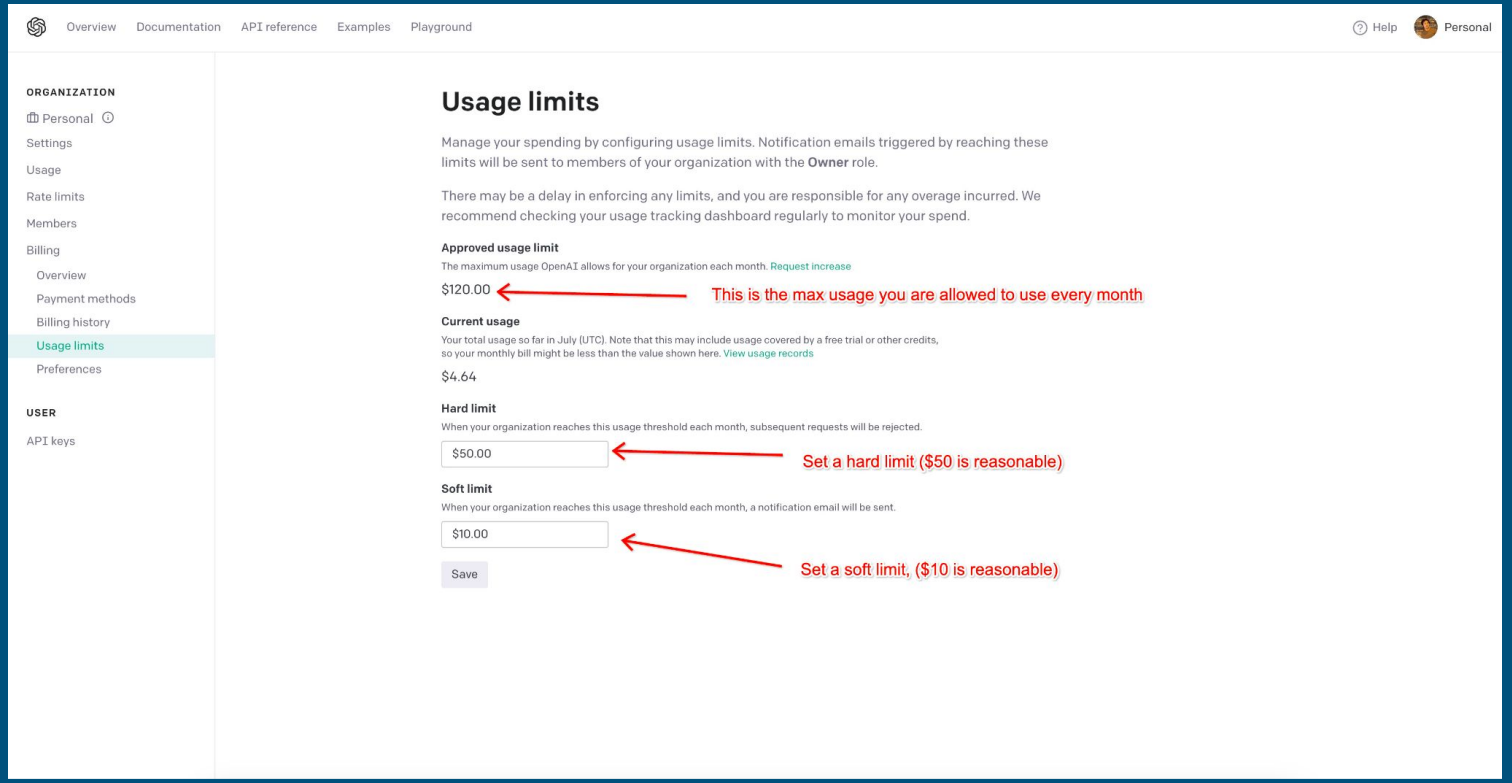
Setting Usage Limits



Click Billing



Click Usage Limits



Set a soft and hard limit