Changelog

Keep track of changes to the OpenAI API. You can also track changes via our public OpenAPI specification which is used to generate our SDKs, documentation, and more. This changelog is maintained in a best effort fashion and may not reflect all changes being made.

May 7th, 2024

- Added support for fine-tuned models to the Batch API.

May 6th, 2024

- Added `stream_options: {"include_usage": true}` parameter to the Chat Completions and Completions APIs. Setting this gives developers access to usage stats when using streaming.

May 2nd, 2024

- Added a new endpoint to delete a message from a thread in the Assistants API.

Apr 29th, 2024

- Added a new function calling option `tool_choice: "required"` to the Chat Completions and Assistants APIs.
- Added a guide for the Batch API and Batch API support for embeddings models

Apr 17th, 2024

- Introduced a series of updates to the Assistants API, including a new file search tool allowing up to 10,000 files per assistant, new token controls, and support for tool choice.

Apr 16th, 2024

- Introduced project based hierarchy for organizing work by projects, including the ability to create API keys and manage rate and cost limits on a per-project basis (cost limits available only for Enterprise customers).

Apr 15th, 2024

- Released Batch API

Apr 9th, 2024

- Released GPT-4 Turbo with Vision in general availability in the API

Apr 4th, 2024

- Added support for seed in the fine-tuning API
- Added support for checkpoints in the fine-tuning API
- Added support for adding Messages when creating a Run in the Assistants API

Apr 1st, 2024

- Added support for filtering Messages by run_id in the Assistants API

Mar 29th, 2024

- Added support for temperature and assistant message creation in the Assistants API

Mar 14th, 2024

- Added support for streaming in the Assistants API

Feb 9th, 2024

- Added `timestamp_granularities` parameter to the Audio API

Feb 1st, 2024

- Released gpt-3.5-turbo-0125, an updated GPT-3.5 Turbo model

Jan 25th, 2024

- Released embedding V3 models and an updated GPT-4 Turbo preview
- Added `dimensions` parameter to the Embeddings API

Dec 20th, 2023

- Added `additional_instructions` parameter to run creation in the Assistants API

Dec 15th, 2023

- Added `logprobs` and `top_logprobs` parameters to the Chat Completions API

Dec 14th, 2023

- Changed function parameters argument on a tool call to be optional

Nov 30th, 2023

- Released OpenAI Deno SDK

Nov 6th, 2023

- Released GPT-4 Turbo Preview, updated GPT-3.5 Turbo, GPT-4 Turbo with Vision, Assistants API, DALL·E 3 in the API, and text-to-speech API
- Deprecated the Chat Completions `functions` parameter in favor of `tools`
- Released OpenAI Python SDK V1.0

Oct 16th, 2023

- Added `encoding_format` parameter to the Embeddings API
- Added `max_tokens` to the Moderation models

Oct 6th, 2023

- Added function calling support to the Fine-tuning API