

Deprecations

Overview

As we launch safer and more capable models, we regularly retire older models. Software relying on OpenAI models may need occasional updates to keep working. Impacted customers will always be notified by email and in our documentation along with [blog posts](#) for larger changes.

This page lists all API deprecations, along with recommended replacements.

Deprecation vs Legacy

We use the term "deprecation" to refer to the process of retiring a model or endpoint. When we announce that a model or endpoint is being deprecated, it immediately becomes deprecated. All deprecated models and endpoints will also have a shut down date. At the time of the shut down, the model or endpoint will no longer be accessible.

We use the term "legacy" to refer to models and endpoints that will no longer receive updates. We tag endpoints and models as legacy to signal to developers where we are moving as a platform and that they should likely migrate to newer models or endpoints. You can expect that a legacy model or endpoint will be deprecated at some point in the future.

Incremental model updates

As [announced](#) in March 2023, we regularly release new versions of `gpt-4` and `gpt-3.5-turbo`.

Each model version is dated with an `-MMDD` suffix; e.g., `gpt-4-0613`. The undated model name, e.g., `gpt-4`, will typically point to the latest version (e.g. `gpt-4` points to `gpt-4-0613`). Users of undated model names will be notified by email typically 2 weeks before any change takes place.

After a new version is launched, older versions will typically be deprecated 3 months later.

Migrating to replacements

Once a model is deprecated, be sure to migrate all usage to a suitable replacement before the shutdown date. Requests to models past the shutdown date will fail.

To help measure the performance of replacement models on your tasks, we have open-sourced [Evals](#), a Python framework for evaluating language models.

If new models perform worse on your tasks, let us know by submitting a [pull request](#) to our Evals repo with examples of the task.

Deprecation history

All deprecations are listed below, with the most recent announcements at the top.

2023-11-06: Chat model updates

On November 6th, 2023, we [announced](#) the release of an updated GPT-3.5-Turbo model (which now comes by default with 16k context) along with deprecation of `gpt-3.5-turbo-0613` and `gpt-3.5-turbo-16k-0613`.

SHUTDOWN DATE	LEGACY MODEL	LEGACY MODEL PRICE	RECOMMENDED REPLACEMENT
2024-06-13	<code>gpt-3.5-turbo-0613</code>	\$0.0015 / 1K input tokens + \$0.0020 / 1K output tokens	<code>gpt-3.5-turbo-1106</code>
2024-06-13	<code>gpt-3.5-turbo-16k-0613</code>	\$0.0030 / 1K input tokens + \$0.0040 / 1K output tokens	<code>gpt-3.5-turbo-1106</code>

Fine-tuned models created from these base models are not effected by this deprecation, but you will no longer be able to create new fine-tuned versions with these models.

2023-08-22: Fine-tunes endpoint

On August 22nd, 2023, we [announced](#) the new fine-tuning API (`/v1/fine_tuning/jobs`) and that the original `/v1/fine-tunes` API along with legacy models (including those fine-tuned with the `/v1/fine-tunes` API) will be shut down on January 04, 2024. This means that models fine-tuned using the `/v1/fine-tunes` API will no longer be accessible and you would have to fine-tune new models with the updated endpoint and associated base models.

Fine-tunes endpoint

SHUTDOWN DATE	SYSTEM	RECOMMENDED REPLACEMENT
2024-01-04	<code>/v1/fine-tunes</code>	<code>/v1/fine_tuning/jobs</code>

2023-07-06: GPT and embeddings

On July 06, 2023, we [announced](#) the upcoming retirements of older GPT-3 and GPT-3.5 models served via the completions endpoint. We also announced the upcoming retirement of our first-generation text embedding models. They will be shut down on January 04, 2024.

InstructGPT models

SHUTDOWN DATE	LEGACY MODEL	LEGACY MODEL PRICE	RECOMMENDED REPLACEMENT
2024-01-04	text-ada-001	\$0.0004 / 1K tokens	gpt-3.5-turbo-instruct
2024-01-04	text-babbage-001	\$0.0005 / 1K tokens	gpt-3.5-turbo-instruct
2024-01-04	text-curie-001	\$0.0020 / 1K tokens	gpt-3.5-turbo-instruct
2024-01-04	text-davinci-001	\$0.0200 / 1K tokens	gpt-3.5-turbo-instruct
2024-01-04	text-davinci-002	\$0.0200 / 1K tokens	gpt-3.5-turbo-instruct
2024-01-04	text-davinci-003	\$0.0200 / 1K tokens	gpt-3.5-turbo-instruct

Pricing for the replacement `gpt-3.5-turbo-instruct` model can be found on the [pricing page](#).

Base GPT models

SHUTDOWN DATE	LEGACY MODEL	LEGACY MODEL PRICE	RECOMMENDED REPLACEMENT
2024-01-04	ada	\$0.0004 / 1K tokens	babbage-002
2024-01-04	babbage	\$0.0005 / 1K tokens	babbage-002
2024-01-04	curie	\$0.0020 / 1K tokens	davinci-002
2024-01-04	davinci	\$0.0200 / 1K tokens	davinci-002
2024-01-04	code-davinci-002	---	gpt-3.5-turbo-instruct

Pricing for the replacement `babbage-002` and `davinci-002` models can be found on the [pricing page](#).

Edit models & endpoint

SHUTDOWN DATE	MODEL / SYSTEM	RECOMMENDED REPLACEMENT
2024-01-04	text-davinci-edit-001	gpt-4
2024-01-04	code-davinci-edit-001	gpt-4

SHUTDOWN DATE	MODEL / SYSTEM	RECOMMENDED REPLACEMENT
2024-01-04	/v1/edits	/v1/chat/completions

Fine-tuning GPT models

SHUTDOWN DATE	LEGACY MODEL	TRAINING PRICE	USAGE PRICE	RECOMMENDED REPLACEMENT
2024-01-04	ada	\$0.0004 / 1K tokens	\$0.0016 / 1K tokens	babbage-002
2024-01-04	babbage	\$0.0006 / 1K tokens	\$0.0024 / 1K tokens	babbage-002
2024-01-04	curie	\$0.003 / 1K tokens	\$0.012 / 1K tokens	davinci-002
2024-01-04	davinci	\$0.03 / 1K tokens	\$0.12 / 1K tokens	davinci-002, gpt-3.5-turbo, gpt-4

First-generation text embedding models

SHUTDOWN DATE	LEGACY MODEL	LEGACY MODEL PRICE	RECOMMENDED REPLACEMENT
2024-01-04	text-similarity-ada-001	\$0.004 / 1K tokens	text-embedding-ada-002
2024-01-04	text-search-ada-doc-001	\$0.004 / 1K tokens	text-embedding-ada-002
2024-01-04	text-search-ada-query-001	\$0.004 / 1K tokens	text-embedding-ada-002
2024-01-04	code-search-ada-code-001	\$0.004 / 1K tokens	text-embedding-ada-002
2024-01-04	code-search-ada-text-001	\$0.004 / 1K tokens	text-embedding-ada-002
2024-01-04	text-similarity-babbage-001	\$0.005 / 1K tokens	text-embedding-ada-002
2024-01-04	text-search-babbage-doc-001	\$0.005 / 1K tokens	text-embedding-ada-002
2024-01-04	text-search-babbage-query-001	\$0.005 / 1K tokens	text-embedding-ada-002
2024-01-04	code-search-babbage-code-001	\$0.005 / 1K tokens	text-embedding-ada-002

SHUTDOWN DATE	LEGACY MODEL	LEGACY MODEL PRICE	RECOMMENDED REPLACEMENT
2024-01-04	code-search-babbage-text-001	\$0.005 / 1K tokens	text-embedding-ada-002
2024-01-04	text-similarity-curie-001	\$0.020 / 1K tokens	text-embedding-ada-002
2024-01-04	text-search-curie-doc-001	\$0.020 / 1K tokens	text-embedding-ada-002
2024-01-04	text-search-curie-query-001	\$0.020 / 1K tokens	text-embedding-ada-002
2024-01-04	text-similarity-davinci-001	\$0.200 / 1K tokens	text-embedding-ada-002
2024-01-04	text-search-davinci-doc-001	\$0.200 / 1K tokens	text-embedding-ada-002
2024-01-04	text-search-davinci-query-001	\$0.200 / 1K tokens	text-embedding-ada-002

2023-06-13: Updated chat models

On June 13, 2023, we announced new chat model versions in the [Function calling and other API updates](#) blog post. The three original versions will be retired in June 2024 at the earliest. As of 01/10/2024, only existing users of this model will be able to continue using this model.

SHUTDOWN DATE	LEGACY MODEL	LEGACY MODEL PRICE	RECOMMENDED REPLACEMENT
at earliest 2024-06-13	gpt-3.5-turbo-0301	\$0.0015 / 1K input tokens + \$0.0020 / 1K output tokens	gpt-3.5-turbo-0613
at earliest 2024-06-13	gpt-4-0314	\$0.03 / 1K input tokens + \$0.06 / 1K output tokens	gpt-4-0613
at earliest 2024-06-13	gpt-4-32k-0314	\$0.06 / 1K input tokens + \$0.12 / 1K output tokens	gpt-4-32k-0613

2023-03-20: Codex models

SHUTDOWN DATE	LEGACY MODEL	RECOMMENDED REPLACEMENT
2023-03-23	code-davinci-002	gpt-4
2023-03-23	code-davinci-001	gpt-4

SHUTDOWN DATE	LEGACY MODEL	RECOMMENDED REPLACEMENT
2023-03-23	code-cushman-002	gpt-4
2023-03-23	code-cushman-001	gpt-4

2022-06-03: Legacy endpoints

SHUTDOWN DATE	SYSTEM	RECOMMENDED REPLACEMENT
2022-12-03	/v1/engines	/v1/models
2022-12-03	/v1/search	View transition guide
2022-12-03	/v1/classifications	View transition guide
2022-12-03	/v1/answers	View transition guide