Document officiel d'OpenAI, copie du 18/11/2026. Source : https://developers.openai.com/api/docs/guides/data-residency

Data residency

For the complete documentation index, see llms.txt. Markdown versions of documentation pages are available by appending `.md` to the page URL.

<span id="data-residency-controls" className="scroll-mt-[var(--docs-scroll-offset)]"

Data residency is a project configuration option that lets you choose the location of infrastructure OpenAI uses to provide services.

Eligible organizations can enable data residency when creating a project in the API Platform. Data residency endpoints are charged a 10% uplift for models released on or after March 5, 2026, that are eligible for data residency.

How does data residency work?

You can select an available region when creating a project.

- **Storage:** For supported endpoints, models, and snapshots, your Customer Content (as defined in your services agreement) is stored at rest in the selected region when the endpoint requires persistence, such as `/v1/batches`.
- **Processing:** If the region supports regional processing, inference for your Customer Content also runs in that region.
- **System data:** Data residency does not apply to system data, which may be stored and processed outside the selected region. This is account data, metadata, and usage data without Customer Content, collected to manage and operate the services. Examples include account information or profiles of end users who directly access the services, analytics, usage statistics, billing information, support requests, and structured output schema.

Sub-processors and regional request processing

OpenAI uses sub-processors to provide its services. For requests sent to `us.api.openai.com`, `eu.api.openai.com`, `au.api.openai.com`, `in.api.openai.com`, or `jp.api.openai.com`, OpenAI uses Cloudflare Regional Services so that TLS termination and HTTPS decryption occur within the selected processing region.

Limitations

Data residency does not apply to:

1. Any transmission or storage of Customer Content outside of the selected region caused by the location of an End User or Customer's infrastructure when accessing the services.
2. Products, services, or content offered by parties other than OpenAI through the Services.
3. Any data other than Customer Content, such as system data.

If your selected Region does not support regional processing, as identified below, OpenAI may also process and temporarily store Customer Content outside of the Region to deliver the services.

How to use data residency

Data residency is configured per-project within your API Organization. As an organization owner:

1. Open the Projects page in the API Platform and select **Create**.
2. Enter a project name and, under **Data and inference residency**, select an available region.
3. Select **Create**, then create an API key for the new project.
4. Use that project's API key with the regional domain prefix listed in the table below, such as `eu.api.openai.com` for Europe.

The project's residency region can't be changed after creation. To select a different region, create a new project.

Select a processing region per request

To enable per-request data residency for a new Global project, follow the creation steps above, leave **Data and inference residency** set to **Global**, and select **Enable per-request data residency** before selecting **Create**. The project remains Global.

Use an API key from that Global project with a regional domain prefix for each request that needs regional processing. The selected endpoint and model must support regional processing, as shown in the table below.

If your existing Global project already supports per-request data residency, you can continue using its API keys.

The following example reuses one client and an API key from a Global project for global, US, and EU requests:

```
from openai import OpenAI

client = OpenAI()

# No processing constraint.
response = client.responses.create(
    model="gpt-6.1-terra",
    input="Reply with OK.",
)
print(response.output_text)

# US processing and storage.
response = client.with_options(
    base_url="https://us.api.openai.com/v1",
).responses.create(
    model="gpt-6.1-terra",
    input="Reply with OK.",
)
print(response.output_text)
```

```
# EU processing and storage.
response = client.with_options(
    base_url="https://eu.api.openai.com/v1",
).responses.create(
    model="gpt-6.1-terra",
    input="Reply with OK.",
)
print(response.output_text)
```

```
require "openai"

client = OpenAI::Client.new

response = client.responses.create(
  model: "gpt-6.1-terra",
  input: "Reply with OK."
)
puts(response.output_text)

response = client.with_options(data_residency: :us).responses.create(
  model: "gpt-6.1-terra",
  input: "Reply with OK."
)
puts(response.output_text)

# EU processing requires a key from an EU-region project.
response = client.with_options(data_residency: :eu, api_key: ENV.fetch("OPENAI_EU_API_KEY"), organization: nil, project: nil).responses.create(
  model: "gpt-6.1-terra",
  input: "Reply with OK."
)
puts(response.output_text)
```

Which models and features are eligible for data residency?

The following models and API services are eligible for data residency today for the regions specified below.

Use **Support by region** to compare regional capabilities and expand the services available in each region. Use **API Endpoint, tool and model support** for complete model lists and a detailed service view. Support for regional storage does not imply support for regional processing.

In Australia, India, and Japan, `gpt-6-luna` supports regional processing. These regions support Standard pay-as-you-go only. Use a project in the selected region. The Responses and Chat Completions APIs support text and image inputs and text output. Hosted tools are unavailable for regional processing in these three regions.

GPT-6.2 Astra and GPT-6.2 Sol support US and EU data residency with Standard, Fast, and Batch processing. Both models support Ultrafast with US data residency only.

Fast mode supports EU data residency for GPT-6.1 Sol, GPT-6.1 Terra, and GPT-6 Luna. Fast mode is not available with EU data residency for GPT-6 Astra or GPT-6 Sol. GPT-6.1 Sol, GPT-6.1 Terra, GPT-6 Sol, and GPT-6 Luna support EU data residency with Standard, Flex, and Batch processing. Ultrafast mode for GPT-6.1 Sol supports US and EU data residency and global processing. GPT-6 Astra Ultrafast supports US data residency and global processing only.

Support by region

The complete, unfiltered regional support table follows. Model snapshots for each service are listed in **API Endpoint, tool and model support**. When regional processing supports only a subset of snapshots, that subset is included in the processing-services cell.

Region	Domain prefix	Regional storage	Regional processing	Additional access requirements	Supported modes	Storage services	Processing services
United States	us.api.openai.com	Yes	Yes	—	Text, Audio, Voice, Image	/v1/audio/transcriptions, /v1/audio/translations, /v1/audio/speech /v1/batches /v1/chat/completions /v1/decisions /v1/embeddings /v1/evals /v1/files /v1/fine_tuning/jobs /v1/images/edits /v1/images/generations /v1/moderations /v1/live/sessions /v1/realtime /v1/realtime/transcription_sessions /v1/realtime/translations /v1/responses /v1/responses File Search /v1/responses Web Search /v1/vector_stores Code Interpreter tool File Search File Uploads Remote MCP server tool Scale Tier Structured Outputs (excluding schema) Supported input modalities	/v1/audio/transcriptions, /v1/audio/translations, /v1/audio/speech /v1/batches /v1/chat/completions /v1/decisions /v1/embeddings /v1/evals /v1/fine_tuning/jobs /v1/images/edits /v1/images/generations /v1/moderations /v1/live/sessions /v1/realtime /v1/realtime/transcription_sessions /v1/realtime/translations /v1/responses /v1/responses File Search /v1/responses Web Search Code Interpreter tool File Search Remote MCP server tool Scale Tier Structured Outputs (excluding schema) Supported input modalities

Region	Domain prefix	Regional storage	Regional processing	Additional access requirements	Supported modes	Storage services	Processing services
Europe (EEA + Switzerland)	eu.api.openai.com	Yes	Yes	—	Text, Audio, Voice, Image*	/v1/audio/transcriptions, /v1/audio/translations, /v1/audio/speech /v1/batches /v1/chat/completions /v1/decisions /v1/embeddings /v1/evals /v1/files /v1/fine_tuning/jobs /v1/images/edits /v1/images/generations /v1/moderations /v1/live/sessions /v1/realtime /v1/realtime/transcription_sessions /v1/realtime/translations /v1/responses /v1/responses File Search /v1/responses Web Search /v1/vector_stores Code Interpreter tool File Search File Uploads Remote MCP server tool Scale Tier Structured Outputs (excluding schema) Supported input modalities	/v1/audio/transcriptions, /v1/audio/translations, /v1/audio/speech /v1/batches /v1/chat/completions /v1/decisions /v1/embeddings /v1/evals /v1/fine_tuning/jobs /v1/images/edits /v1/images/generations /v1/moderations /v1/live/sessions /v1/realtime /v1/realtime/transcription_sessions /v1/realtime/translations /v1/responses /v1/responses File Search /v1/responses Web Search Code Interpreter tool File Search Remote MCP server tool Scale Tier Structured Outputs (excluding schema) Supported input modalities
Australia*	au.api.openai.com	Yes	Yes	—	Text, Audio, Voice, Image	/v1/audio/transcriptions, /v1/audio/translations, /v1/audio/speech /v1/batches /v1/chat/completions /v1/decisions /v1/embeddings /v1/files /v1/fine_tuning/jobs /v1/images/edits /v1/images/generations /v1/moderations /v1/responses /v1/responses File Search /v1/responses Web Search /v1/vector_stores Code Interpreter tool File Search File Uploads Remote MCP server tool Scale Tier Structured Outputs (excluding schema) Supported input modalities	/v1/chat/completions (gpt-6-luna) /v1/responses (gpt-6-luna)
Canada*	ca.api.openai.com	Yes	No	—	Text, Audio, Voice, Image	/v1/audio/transcriptions, /v1/audio/translations, /v1/audio/speech /v1/batches /v1/chat/completions /v1/decisions /v1/embeddings /v1/files /v1/fine_tuning/jobs /v1/images/edits /v1/images/generations /v1/moderations /v1/responses /v1/responses File Search /v1/responses Web Search /v1/vector_stores Code Interpreter tool File Search File Uploads Remote MCP server tool Scale Tier Structured Outputs (excluding schema) Supported input modalities	None
Japan*	jp.api.openai.com	Yes	Yes	—	Text, Audio, Voice, Image	/v1/audio/transcriptions, /v1/audio/translations, /v1/audio/speech /v1/batches /v1/chat/completions /v1/decisions	/v1/chat/completions (gpt-6-luna) /v1/responses (gpt-6-luna)

Region	Domain prefix	Regional storage	Regional processing	Additional access requirements	Supported modes	Storage services	Processing services
						/v1/embeddings /v1/files /v1/fine_tuning/jobs /v1/images/edits /v1/images/generations /v1/moderations /v1/responses /v1/responses File Search /v1/responses Web Search /v1/vector_stores Code Interpreter tool File Search File Uploads Remote MCP server tool Scale Tier Structured Outputs (excluding schema) Supported input modalities	
India*	in.api.openai.com	Yes	Yes	—	Text, Audio, Voice, Image	/v1/audio/transcriptions, /v1/audio/translations, /v1/audio/speech /v1/batches /v1/chat/completions /v1/decisions /v1/embeddings /v1/files /v1/fine_tuning/jobs /v1/images/edits /v1/images/generations /v1/moderations /v1/responses /v1/responses File Search /v1/responses Web Search /v1/vector_stores Code Interpreter tool File Search File Uploads Remote MCP server tool Scale Tier Structured Outputs (excluding schema) Supported input modalities	/v1/chat/completions (gpt-6-luna) /v1/responses (gpt-6-luna)
Singapore*	sg.api.openai.com	Yes	No	—	Text, Audio, Voice, Image	/v1/audio/transcriptions, /v1/audio/translations, /v1/audio/speech /v1/batches /v1/chat/completions /v1/decisions /v1/embeddings /v1/files /v1/fine_tuning/jobs /v1/images/edits /v1/images/generations /v1/moderations /v1/responses /v1/responses File Search /v1/responses Web Search /v1/vector_stores Code Interpreter tool File Search File Uploads Remote MCP server tool Scale Tier Structured Outputs (excluding schema) Supported input modalities	None
South Korea*	kr.api.openai.com	Yes	No	—	Text, Audio, Voice, Image	/v1/audio/transcriptions, /v1/audio/translations, /v1/audio/speech /v1/batches /v1/chat/completions /v1/decisions /v1/embeddings /v1/files /v1/fine_tuning/jobs /v1/images/edits /v1/images/generations /v1/moderations /v1/responses /v1/responses File Search /v1/responses Web Search /v1/vector_stores Code Interpreter tool	None

Region	Domain prefix	Regional storage	Regional processing	Additional access requirements	Supported modes	Storage services	Processing services
						File Search File Uploads Remote MCP server tool Scale Tier Structured Outputs (excluding schema) Supported input modalities	
United Kingdom*	gb.api.openai.com	Yes	No	—	Text, Audio, Voice, Image	/v1/audio/transcriptions, /v1/audio/translations, /v1/audio/speech /v1/batches /v1/chat/completions /v1/decisions /v1/embeddings /v1/files /v1/fine_tuning/jobs /v1/images/edits /v1/images/generations /v1/moderations /v1/responses /v1/responses File Search /v1/responses Web Search /v1/vector_stores Code Interpreter tool File Search File Uploads Remote MCP server tool Scale Tier Structured Outputs (excluding schema) Supported input modalities	None
United Arab Emirates*	ae.api.openai.com	Yes	Yes	Contact sales to access this region	Text, Audio, Voice, Image	/v1/audio/transcriptions, /v1/audio/translations, /v1/audio/speech /v1/batches /v1/chat/completions /v1/decisions /v1/embeddings /v1/files /v1/fine_tuning/jobs /v1/images/edits /v1/images/generations /v1/moderations /v1/responses /v1/responses File Search /v1/responses Web Search /v1/vector_stores Code Interpreter tool File Search File Uploads Remote MCP server tool Scale Tier Structured Outputs (excluding schema) Supported input modalities	/v1/chat/completions (gpt-5.6-luna, gpt-5.5-2026-04-23, gpt-5.2-2025-12-11) /v1/embeddings (text-embedding-3-large) /v1/responses (gpt-5.5-pro-2026-04-23, gpt-5.6-luna, gpt-5.5-2026-04-23, gpt-5.2-2025-12-11)

* Image support in these regions requires additional approval. Contact sales.

API Endpoint, tool and model support

Endpoint or feature	Service	Storage regions	Processing regions	Supported models and snapshots	Regional processing snapshot exceptions	Notes
/v1/audio/transcriptions, /v1/audio/translations, /v1/audio/speech	Audio	All listed regions	United States, Europe (EEA + Switzerland)	tts-1, whisper-1, gpt-4o-tts, gpt-4o-transcribe, gpt-4o-mini-transcribe, gpt-transcribe	None	—
/v1/batches	Batches	All listed regions	United States, Europe (EEA + Switzerland)	gpt-6.2-astra, gpt-6.2-sol, gpt-6-astra, gpt-6.1-sol, gpt-6.1-terra, gpt-6-sol, gpt-6-luna, gpt-5.5-pro-2026-04-23, gpt-5.4-pro-2026-03-05, gpt-5.2-pro-2025-12-11, gpt-5-pro-2025-10-06, gpt-5.6-sol, gpt-5.6-terra, gpt-5.6-luna, gpt-5.5-2026-04-23, gpt-5.4-2026-03-05, gpt-5-2025-08-07, gpt-5.4-mini-2026-03-17, gpt-5.4-nano-2026-03-17, gpt-5.2-2025-12-11, gpt-5.1-2025-11-13, gpt-5-mini-2025-08-07, gpt-5-nano-2025-08-07, gpt-4.1-2025-04-14, gpt-4.1-mini-2025-04-14, gpt-4.1-nano-2025-04-14, o3-	None	GPT-6.2 Astra and GPT-6.2 Sol support US and EU data residency with Batch processing. GPT-6.1 Sol, GPT-6.1 Terra, GPT-6 Sol, and GPT-6 Luna support EU data residency with Standard, Flex, and Batch processing. GPT-6.1 Sol and GPT-6.1 Terra support only US and EU data residency.

Endpoint or feature	Service	Storage regions	Processing regions	Supported models and snapshots	Regional processing snapshot exceptions	Notes
				2025-04-16, o4-mini-2025-04-16, o1-pro, o1-pro-2025-03-19, o3-mini-2025-01-31, o1-2024-12-17, gpt-4o-2024-11-20, gpt-4o-2024-08-06, gpt-4o-mini-2024-07-18, gpt-4-turbo-2024-04-09, gpt-4-0613, gpt-3.5-turbo-0125		
/v1/chat/completions	Chat Completions	All listed regions	United States, Europe (EEA + Switzerland), United Arab Emirates, Australia, India, Japan	gpt-6.2-astra, gpt-6.2-sol, gpt-6-astra, gpt-6.1-sol, gpt-6.1-terra, gpt-6-sol, gpt-6-luna, gpt-5.6-sol, gpt-5.6-terra, gpt-5.6-luna, gpt-5.5-2026-04-23, gpt-5.4-2026-03-05, gpt-5.4-mini-2026-03-17, gpt-5.4-nano-2026-03-17, gpt-5.2-2025-12-11, gpt-5.1-2025-11-13, gpt-5-2025-08-07, gpt-5-mini-2025-08-07, gpt-5-nano-2025-08-07, gpt-4.1-2025-04-14, gpt-4.1-mini-2025-04-14, gpt-4.1-nano-2025-04-14, o3-mini-2025-01-31, o3-2025-04-16, o4-mini-2025-04-16, o1-2024-12-17, gpt-4o-2024-11-20, gpt-4o-2024-08-06, gpt-4o-mini-2024-07-18, gpt-4-turbo-2024-04-09, gpt-4-0613, gpt-3.5-turbo-0125	United Arab Emirates: gpt-5.6-luna, gpt-5.5-2026-04-23, gpt-5.2-2025-12-11 Australia, India, Japan: gpt-6-luna	GPT-6.2 Astra and GPT-6.2 Sol support US and EU data residency with Standard and Fast processing. Fast mode supports EU data residency for GPT-6.1 Sol, GPT-6.1 Terra, and GPT-6 Luna. Fast mode is not available with EU data residency for GPT-6 Astra or GPT-6 Sol. GPT-6.1 Sol and GPT-6.1 Terra support only US and EU data residency.
/v1/decisions	Decisions	All listed regions	United States, Europe (EEA + Switzerland)	gpt-6-luna	None	Available in all supported API regions. Regional processing is supported in the United States and Europe (EEA + Switzerland). Prompt caching is subject to the retention limitations below.
/v1/embeddings	Embeddings	All listed regions	United States, Europe (EEA + Switzerland), United Arab Emirates	text-embedding-3-small, text-embedding-3-large, text-embedding-ada-002	United Arab Emirates: text-embedding-3-large	—
/v1/evals	Evals	United States, Europe (EEA + Switzerland)	United States, Europe (EEA + Switzerland)	Service-level support	None	—
/v1/files	Files	All listed regions	None	Service-level support	None	—
/v1/fine_tuning/jobs	Fine-tuning	All listed regions	United States, Europe (EEA + Switzerland)	gpt-4o-2024-08-06, gpt-4o-mini-2024-07-18, gpt-4.1-2025-04-14, gpt-4.1-mini-2025-04-14	None	—
/v1/images/edits	Images	All listed regions	United States, Europe (EEA + Switzerland)	gpt-image-2.5-sunburst, gpt-image-2.5-sunburst-2026-09-08, gpt-image-2.5-flare, gpt-image-2.5-flare-2026-09-08, gpt-image-2, gpt-image-1.5, gpt-image-1-mini	None	—
/v1/images/generations	Images	All listed regions	United States, Europe (EEA + Switzerland)	gpt-image-2.5-sunburst, gpt-image-2.5-sunburst-2026-09-08, gpt-image-2.5-flare, gpt-image-2.5-flare-2026-09-08, gpt-image-2, gpt-image-1.5, gpt-image-1-mini	None	—
/v1/moderations	Moderation	All listed regions	United States, Europe (EEA + Switzerland)	omni-moderation-latest	None	—
/v1/moderations (CSAM classifier)	Moderation	Not specified	Not specified	omni-moderation-csam-latest	—	Available to approved organizations. Regional processing coverage is not specified for this model.
/v1/live/sessions	GPT-Live	United States, Europe (EEA + Switzerland)	United States, Europe (EEA + Switzerland)	gpt-live-1	None	—
/v1/realtime	Realtime	United States, Europe (EEA + Switzerland)	United States, Europe (EEA + Switzerland)	gpt-realtime, gpt-realtime-1.5, gpt-realtime-mini, gpt-realtime-2, gpt-realtime-2.1, gpt-realtime-2.1-mini	None	—
/v1/realtime/transcription_sessions	Realtime	United States, Europe (EEA	United States, Europe (EEA + Switzerland)	gpt-realtime-whisper, gpt-live-transcribe, gpt-transcribe	None	—

Endpoint or feature	Service	Storage regions	Processing regions	Supported models and snapshots	Regional processing snapshot exceptions	Notes
		+ Switzerland)				
`/v1/realtime/translations`	Realtime	United States, Europe (EEA + Switzerland)	United States, Europe (EEA + Switzerland)	`gpt-realtime-translate`	None	—
`/v1/responses`	Responses	All listed regions	United States, Europe (EEA + Switzerland), United Arab Emirates, Australia, India, Japan	`gpt-6.2-astra`, `gpt-6.2-sol`, `gpt-6-astra`, `gpt-6.1-sol`, `gpt-6.1-terra`, `gpt-6-sol`, `gpt-6-luna`, `gpt-5.5-pro-2026-04-23`, `gpt-5.4-pro-2026-03-05`, `gpt-5.2-pro-2025-12-11`, `gpt-5-pro-2025-10-06`, `gpt-5.6-sol`, `gpt-5.6-terra`, `gpt-5.6-luna`, `gpt-5.5-2026-04-23`, `gpt-5.4-2026-03-05`, `gpt-5-2025-08-07`, `gpt-5.4-mini-2026-03-17`, `gpt-5.4-nano-2026-03-17`, `gpt-5.2-2025-12-11`, `gpt-5.1-2025-11-13`, `gpt-5-mini-2025-08-07`, `gpt-5-nano-2025-08-07`, `gpt-4.1-2025-04-14`, `gpt-4.1-mini-2025-04-14`, `gpt-4.1-nano-2025-04-14`, `o3-2025-04-16`, `o4-mini-2025-04-16`, `o1-pro`, `o1-pro-2025-03-19`, `o3-mini-2025-01-31`, `o1-2024-12-17`, `gpt-4o-2024-11-20`, `gpt-4o-2024-08-06`, `gpt-4o-mini-2024-07-18`, `gpt-4-turbo-2024-04-09`, `gpt-4-0613`, `gpt-3.5-turbo-0125`	United Arab Emirates: `gpt-5.5-pro-2026-04-23`, `gpt-5.6-luna`, `gpt-5.5-2026-04-23`, `gpt-5.2-2025-12-11` Australia, India, Japan: `gpt-6-luna`	GPT-6.2 Astra and GPT-6.2 Sol support US and EU data residency with Standard and Fast processing. Fast mode supports EU data residency for GPT-6.1 Sol, GPT-6.1 Terra, and GPT-6 Luna. Fast mode is not available with EU data residency for GPT-6 Astra or GPT-6 Sol. GPT-6.1 Sol and GPT-6.1 Terra support only US and EU data residency.
`/v1/responses File Search`	Responses	All listed regions	United States, Europe (EEA + Switzerland)	Service-level support	None	—
`/v1/responses Web Search`	Responses	All listed regions	United States, Europe (EEA + Switzerland)	Service-level support	None	—
`/v1/vector_stores`	Vector stores	All listed regions	None	Service-level support	None	—
`Code Interpreter tool`	Tools	All listed regions	United States, Europe (EEA + Switzerland)	Service-level support	None	—
`File Search`	Tools	All listed regions	United States, Europe (EEA + Switzerland)	Service-level support	None	—
`File Uploads`	Files	All listed regions	None	Service-level support	None	Supported when used with base64 file uploads.
`Remote MCP server tool`	Tools	All listed regions	United States, Europe (EEA + Switzerland)	Service-level support	None	MCP servers are third-party services. Data sent to an MCP server is subject to its data residency policies.
`Scale Tier`	Other	All listed regions	United States, Europe (EEA + Switzerland)	Service-level support	None	—
`Structured Outputs (excluding schema)`	Other	All listed regions	United States, Europe (EEA + Switzerland)	Service-level support	None	—
`Supported input modalities`	Other	All listed regions	United States, Europe (EEA + Switzerland)	`Text`, `Image`, `Audio/Voice`	None	—

Endpoint limitations

<span id="v1decisions-1" className="scroll-mt-[var(--docs-scroll-offset)]"

/v1/decisions

The Decisions API is available in all supported API regions. Regional processing is supported in the United States and Europe (EEA + Switzerland). Availability in a region does not imply that inference is performed in that region.

- Extended prompt caching in regions that do not support Regional processing may require that OpenAI process and temporarily store Customer Content outside of the Region to deliver the services.

<span id="v1chatcompletions-1" className="scroll-mt-[var(--docs-scroll-offset)]"

/v1/chat/completions

- Cannot set store=true in non-US regions.
- Extended prompt caching in regions that do not support Regional processing may require that OpenAI process and temporarily store Customer Content outside of the Region to deliver the services.

<span id="v1responses-1" className="scroll-mt-[var(--docs-scroll-offset)]"

/v1/responses

- Extended prompt caching in regions that do not support Regional processing may require that OpenAI process and temporarily store Customer Content outside of the Region to deliver the services.

/v1/live/sessions

GPT-Live sessions are eligible for Zero Data Retention. With Zero Data Retention enabled, `store` is treated as `false`, even if a request sets it to `true`.

Session storage is disabled by default. For projects with session storage enabled, `store: true` retains the completed session recording for 30 days so that it can be downloaded or used to start a forked session. Stored sessions and their index expire after 30 days. Recording downloads and forking require a data policy that permits persistence and are not available with Zero Data Retention.

Setting `store: false` on a fork prevents storage of the new session; it does not delete the source recording or remove the authorization required to read it. The API does not provide a public stored-session deletion endpoint.

GPT-Live supports data residency in the United States and Europe. Delegated backend models and tools have their own data controls; check the applicable endpoint and feature entries on this page.

/v1/realtime

Tracing is not currently EU data residency compliant for `/v1/realtime`.