Starting April 29, 2025, Gemini 1.5 Pro and Gemini 1.5 Flash models are not available in projects that have no prior usage of these models, including new projects. For details, see <u>Model versions and lifecycle</u>

(/vertex-ai/generative-ai/docs/learn/model-versions#legacy-stable).

# Manage your RAG knowledge base (corpus)

The <u>VPC-SC security control</u> (/vertex-ai/generative-ai/docs/security-controls) is supported by RAG Engine. Data residency, CMEK, and AXT security controls aren't supported.

This page describes how you can manage your corpus for RAG tasks by performing corpus management and file management.

## Corpus management

A corpus, also referred to as an index, is a collection of documents or source of information. The corpus can be queried to retrieve relevant contexts for response generation. When creating a corpus for the first time, the process might take an additional minute.

The following corpus operations are supported:

Operation	Description	Parameters	Examples
Create a	Create a RAG corpus to	Create parameters v1beta1	Create example v1beta1
RAG corpus	s.import or upload	(/vertex-ai/generative-ai/docs/model	- (/vertex-ai/generative-
	documents.	reference/rag-api#create-a-rag-	ai/docs/model-reference/rag-
		corpus-params-api)	api#create-a-rag-corpus-example-api)
		<u>Create parameters v1</u>	<u>Create example v1</u>
		(/vertex-ai/generative-ai/docs/model	- (/vertex-ai/generative-
		reference/rag-api-v1#create-a-rag-	ai/docs/model-reference/rag-api-
		corpus-params-api)	v1#create-a-rag-corpus-example-api)

Operation	Description	Parameters	Examples
Update a RAG corpus	Update a previously- s. created RAG corpus to import or upload documents.	Update parameters v1beta1 (/vertex-ai/generative-ai/docs/model- reference/rag-api#update-a-rag- corpus-params-api) Update parameters v1 (/vertex-ai/generative-ai/docs/model- reference/rag-api-v1#update-a-rag- corpus-params-api)	ai/docs/model-reference/rag- api#update-a-rag-corpus-example- api)
List a RAG corpus.	List all of the RAG corpora.	List parameters v1beta1  (/vertex-ai/generative-ai/docs/model- reference/rag-api#list-rag-corpora- params-api)  List parameters v1  (/vertex-ai/generative-ai/docs/model- reference/rag-api-v1#list-rag-corpora- params-api)	ai/docs/model-reference/rag-api#list- rag-corpora-example-api) <u>List example v1</u> - (/vertex-ai/generative-
Get a RAG corpus.	Get the metadata describing the RAG corpus.	Get parameters v1beta1 (/vertex-ai/generative-ai/docs/model- reference/rag-api#get-a-rag-corpus- params-api) Get parameters v1 (/vertex-ai/generative-ai/docs/model- reference/rag-api-v1#get-a-rag- corpus-params-api)	ai/docs/model-reference/rag-api#get- a-rag-corpus-example-api) Get example v1
Delete a RAG corpus	Delete the RAG corpus.	Delete parameters v1beta1 (/vertex-ai/generative-ai/docs/model- reference/rag-api#delete-a-rag- corpus-params-api) Delete parameters v1 (/vertex-ai/generative-ai/docs/model- reference/rag-api-v1#delete-a-rag- corpus-params-api)	ai/docs/model-reference/rag- api#delete-a-rag-corpus-example-api) Delete example v1

Concurrent operations on corpora aren't supported. For more information, see the <u>RAG API reference</u> for v1beta1 (/vertex-ai/generative-ai/docs/model-reference/rag-api) or <u>RAG API reference for v1</u> (/vertex-ai/generative-ai/docs/model-reference/rag-api-v1).

## File management

The following file operations are supported:

Operation	Description	Parameters	Examples
Upload a RAG file.	Upload a file from local storage with additional information that provides context to the LLM to generate more accurate responses.	Upload parameters v1beta1 (/vertex-ai/generative- ai/docs/model-reference/rag- api#upload-a-rag-file-params-api) Upload parameters v1 (/vertex-ai/generative- ai/docs/model-reference/rag-api- v1#upload-a-rag-file-params-api)	Upload example v1beta1  (/vertex-ai/generative- ai/docs/model-reference/rag- api#upload-a-rag-file-example- api)  Upload example v1  (/vertex-ai/generative- ai/docs/model-reference/rag-ap v1#upload-a-rag-file-example-ap
Import RAG files.	Import a set of files from some other storage into a storage location.	Import parameters v1beta1 (/vertex-ai/generative- ai/docs/model-reference/rag- api#import-rag-files-params-api) Import parameters v1 (/vertex-ai/generative- ai/docs/model-reference/rag-api- v1#import-rag-files-params-api)	Import example v1beta1 (/vertex-ai/generative- ai/docs/model-reference/rag- api#import-rag-files-example-api Import example v1 (/vertex-ai/generative- ai/docs/model-reference/rag-api v1#import-rag-files-example-api)
List RAG files.	List all of the files in the RAG corpus.	List parameters v1beta1 (/vertex-ai/generative- ai/docs/model-reference/rag- api#list-rag-files-params-api) List parameters v1 (/vertex-ai/generative- ai/docs/model-reference/rag-api- v1#list-rag-files-params-api)	List example v1beta1 (/vertex-ai/generative-ai/docs/model-reference/rag-api#list-rag-files-example-api) List example v1 (/vertex-ai/generative-ai/docs/model-reference/rag-api v1#list-rag-files-example-api)
Get a RAG file.	Get details about a RAG file for use by the LLM.	Get parameters v1beta1 (/vertex-ai/generative- ai/docs/model-reference/rag- api#get-a-rag-file-params-api) Get parameters v1 (/vertex-ai/generative- ai/docs/model-reference/rag-api- v1#get-a-rag-file-params-api)	Get example v1beta1 (/vertex-ai/generative- ai/docs/model-reference/rag- api#get-a-rag-file-example-api) Get example v1 (/vertex-ai/generative- ai/docs/model-reference/rag-api v1#get-a-rag-file-example-api)
Delete a RAG file.	Delete a file from the RAG corpus.	Delete parameters v1beta1 (/vertex-ai/generative- ai/docs/model-reference/rag- api#delete-a-rag-file-params-api) Delete parameters v1 (/vertex-ai/generative- ai/docs/model-reference/rag-api- v1#delete-a-rag-file-params-api)	Delete example v1beta1 (/vertex-ai/generative- ai/docs/model-reference/rag- api#delete-a-rag-file-example-api Delete example v1 (/vertex-ai/generative- ai/docs/model-reference/rag-api v1#delete-a-rag-file-example-api)

#### For more information, see the <u>RAG API reference for v1beta1</u>

(/vertex-ai/generative-ai/docs/model-reference/rag-api) or <u>RAG API reference for v1</u> (/vertex-ai/generative-ai/docs/model-reference/rag-api-v1).

### What's next

• Vertex Al RAG Engine quotas (/vertex-ai/generative-ai/docs/quotas#rag-quotas)

Except as otherwise noted, the content of this page is licensed under the <u>Creative Commons Attribution 4.0 License</u> (https://creativecommons.org/licenses/by/4.0/), and code samples are licensed under the <u>Apache 2.0 License</u> (https://www.apache.org/licenses/LICENSE-2.0). For details, see the <u>Google Developers Site Policies</u> (https://developers.google.com/site-policies). Java is a registered trademark of Oracle and/or its affiliates.

Last updated 2025-06-06 UTC.