📄 API Documentation

**1. Embed Document**

Endpoint: **/api/embedding**

Method: POST

Description: Uploads and embeds a document (TXT, PDF, or DOCX) into a vector database using Google Gemini embeddings.

✅ Request Parameters (Form Data):

document (file) — Required. The file to be embedded.

**📤 Response:**

Success:

{

"status": "success",

"message": "Document embedded successfully.",

"document\_id": "unique-document-id"

}

Error:

{

"status": "error",

"message": "Failed to embed document.",

"error\_details": "Document content is empty."

🔍 **Query Document API**

Method

POST

Endpoint

**/api/query**

Description

This API allows querying a previously uploaded and embedded document. It retrieves relevant text segments using Retrieval-Augmented Generation (RAG) and provides a response generated from those segments. Ideal for asking specific questions related to document content.

Request Parameters

All parameters are submitted as form-data:

query (required):

The user’s question or prompt that should be answered based on the document content.

conversation\_id (required):

Unique identifier used to maintain conversation or session context.

document\_id (required):

Identifier for the specific document that was previously uploaded and embedded.

require\_citations (optional):

Flag indicating whether the response should include references to the source text. Accepts true or false. Default is false.

**Success Response**

Returns a generated answer relevant to the query and the specified document.

The response is structured in a way that contains only the answer text unless citations are requested.

**Error Handling**

Missing or invalid fields:

Returned if any required parameter is not provided or invalid.

**Invalid document reference:**

Returned if the document\_id does not correspond to any embedded document.

**Internal server error:**

Returned if the system encounters any unexpected error during processing.

**Usage Notes**

This API is part of a pipeline that starts with document upload and embedding. Make sure the document is successfully embedded before querying.

The conversation ID can be reused to maintain chat context across multiple queries.

When require\_citations is enabled, the system includes text references from the source document in the response.