RAG

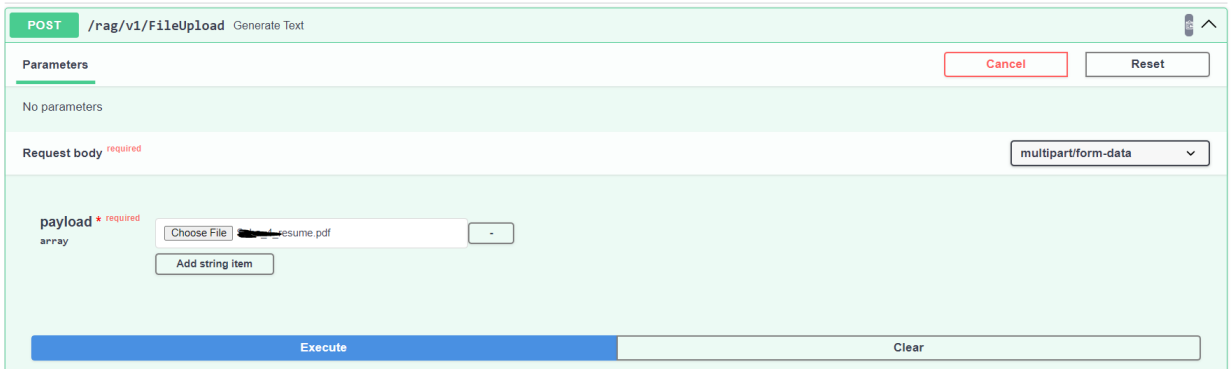
To try the endpoints, click on “try it out”

Endpoints:

1. FileUpload -->

   It is used to upload document with pdf format and return vectorestoreid of vectorestore and blobname for the pdf.

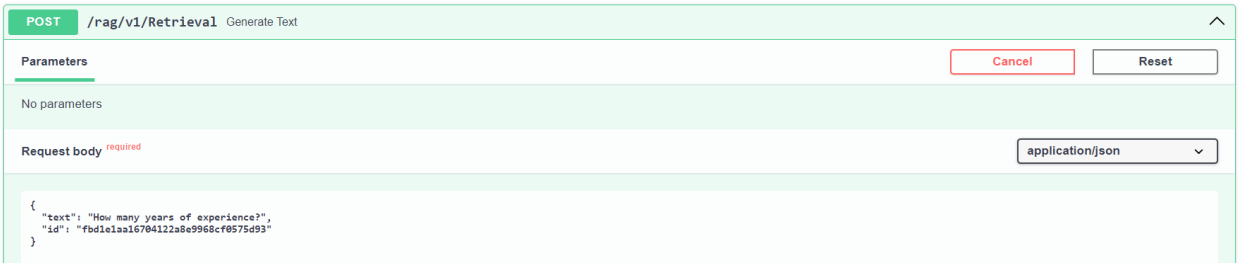
   Input Payload:



2. Retrieval -->

   This will take the id (vectorestoreid returned by FileUpload), and text(Prompt) as Input and will return the rag response by searching the document along with hallucination score and source of document.

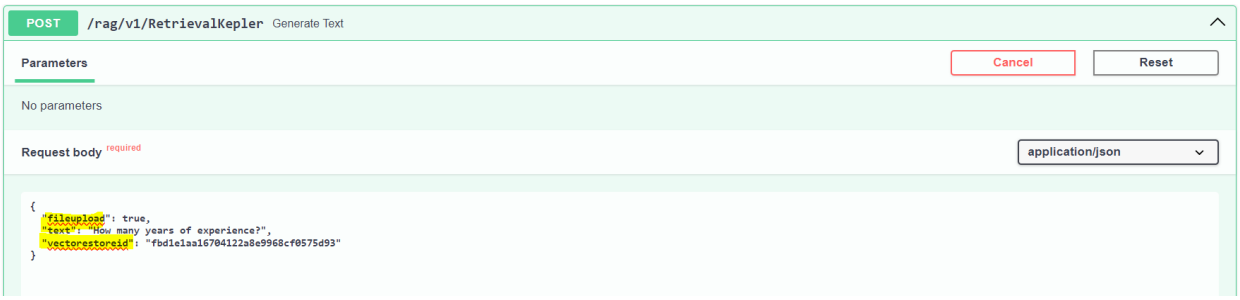
   Input Payload:



3. RetrievalKepler -->

   This will take keys FileUpload(True if using File Upload else False when using Vectorestore Caching), text, vectorestoreid as Input and will return a rag response along with the score.

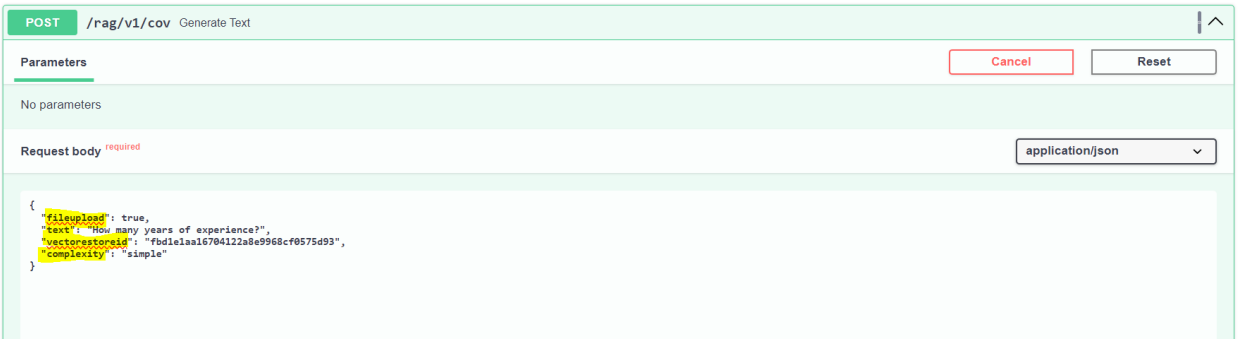
   Input Payload:



4. CoVe – ->

This will take keys FileUpload(True if using File Upload else False when using Vectorestore Caching), text, and vectorestoreid as Input and will return the response with 5 more variants of questions generated by LLM to verify the answer and a refined final response.

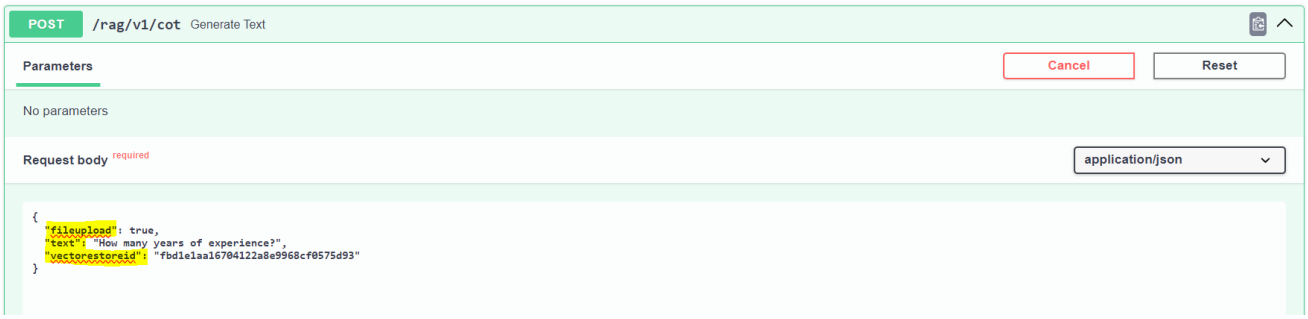
Input Payload :



5. Chain of thought (CoT)-

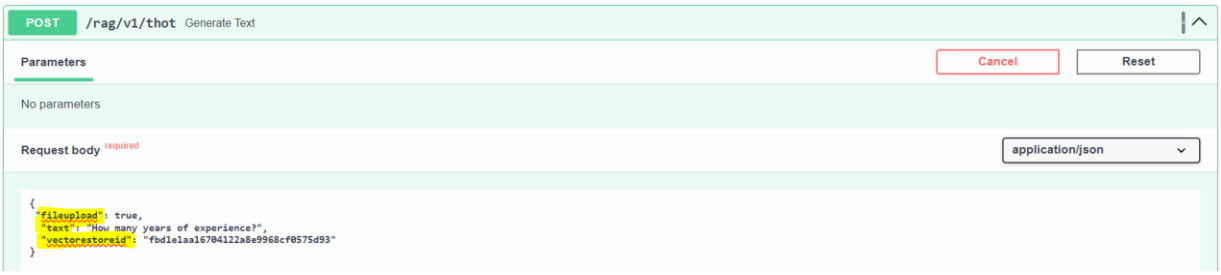
This will take keys FileUpload(True if using File Upload else False when using Vectorestore Caching), text, vectorestoreid as Input and will return the response.
Chain of Thought response provides you explanation steps and reasoning behind and from where (source referred which document)

Input Payload:



6. Thread of Thought(ThoT)-

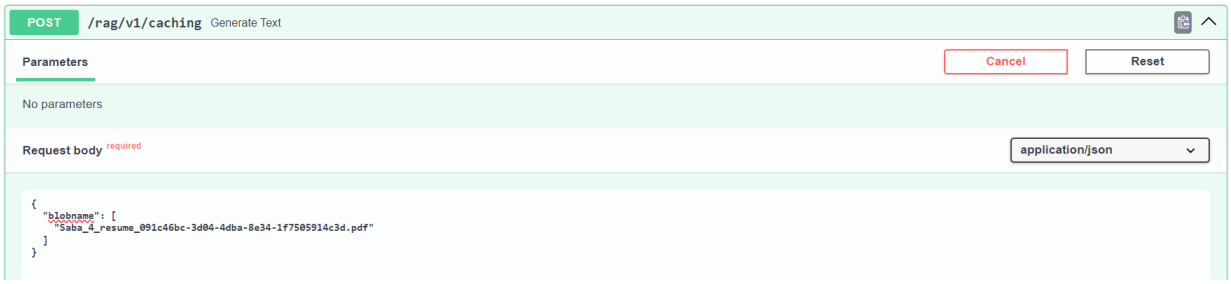
This will take keys FileUpload(True if using File Upload else False when using Vectorestore Caching), text, vectorestoreid as Input and will return Chain of verification response.
Thot is efficient in more descriptive and complex information spread over the file.
Input Payload :



7. Caching-

This will take blobname which is generated while uploading file and return a array of length 2. First element will be the Cache id, Second element will be the Cache id which is removed from the cache if cache is full else 0.

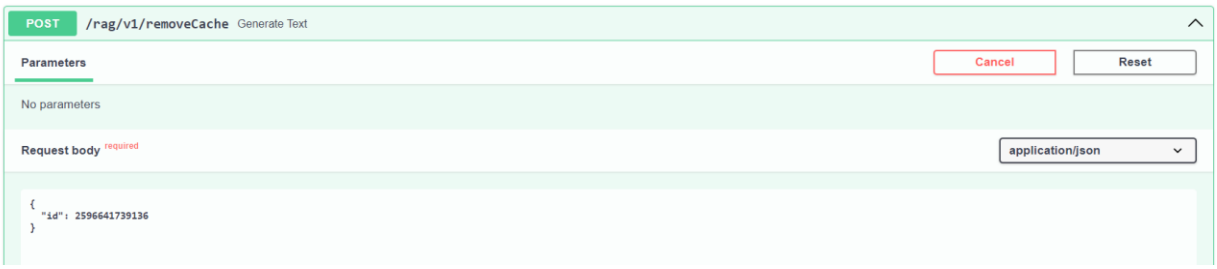
Input Payload:



8. RemoveCache-

It will take input as cache id and will remove the cached Vectorestore from the Cache.

Input Payload:



9. Show_Score –

This will take input as Prompt, response which RetrivalKepler gave, SourceArr which is nothing but page content from RetrievalKepler as Input and Will return Hallucination Score.

Input Payload: