Help Mate AI- RAG Project

Prepared by – Arun Kumar Singh (asingh.iitr@gmail.com) Date: 31-Mar-2024

Goal:

Build a robust generative search system capable of effectively and accurately answering questions from a policy document, by using RAG (Retrieval augmented generation) concept.

Data Source: Policy document present in PDF form as below:

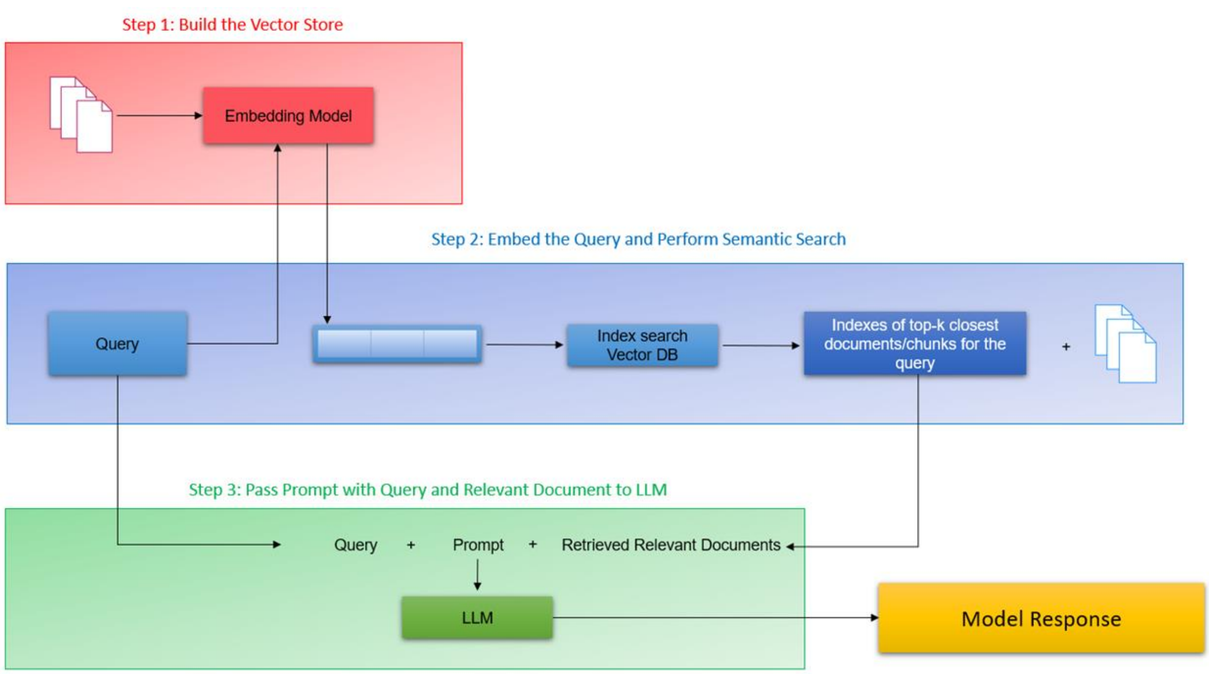
https://github.com/arunksinghbuee/helpmate-ai/blob/main/insurance-document/Principal-Sample-Life-Insurance-Policy.pdf

Collab Notebook PDF: https://github.com/arunksinghbuee/helpmate-ai/blob/main/Mr.HelpMate%20AI%20Project.pdf

System Design:

RAG pipeline mainly consists of 3 layers:

1. Embedding Layer
2. Search and Rank Layer
3. Generation Layer



Embedding Layer

This layer is a crucial component of RAG models, serving as the first layer. It contains an embedding model trained on extensive text and code datasets to learn word and phrase relationships. This layer enables systems to understand text meaning and its semantic relation to queries, crucial for tasks like question answering, summarization, and machine translation. It generates embeddings for text corpora, aiding the RAG model in comprehending queries and generating relevant responses.

Embedding generation code -> Refer https://github.com/arunksinghbuee/helpmate-ai/

There are multiple types of chunking methods which can be used as per nature of data.

Chunking Method	Description	Advantages	Disadvantages
Use entire document without chunking	The entire document is treated as a single chunk.	Simple and easy to implement.	The chunk may not always fit within the context window. This may result in extraneous contexts, thereby affecting the quality of the embedding.
Fixed-size chunking	The document is split into fixed-size windows with each window representing a separate document chunk.	Simple and easy to implement. Having a consistent chunk size will make the system consistent.	May cut off the context in between chunks, resulting in information loss.
Chunking on natural delimiters	Natural delimiters in the document are used to determine the boundaries of each chunk (sentence, paragraph, etc.).	Can result in more meaningful chunks as it utilises the structure of the document and the natural breaks within it.	May be time-consuming to find the right delimiters.
Overlapping chunks	The document is split into fixed-size overlapping windows.	Simple and easy to implement. Redundancy is offset by higher accuracy and latency.	May result in information redundancy across different chunks.

By looking at provided document, most of the similar context data is present on each page and it is further divided into sections & paragraphs on each page.

Word length in each page varies up to maximum of 418 words which is suitable for creation of embedding using **text-embedding-ada-002** OpenAI embedding model.

```python
insurance_pdf_data['Text_Length']
```

```
0      30
1       5
2     230
3       5
4     110
     ...
59    285
60    418
61    322
62      5
63      8
Name: Text_Length, Length: 64, dtype: int64
```

Embedding models

OpenAI offers two powerful third-generation embedding model (denoted by `-3` in the model ID). You can read the embedding v3 announcement blog post for more details.

Usage is priced per input token, below is an example of pricing pages of text per US dollar (assuming ~800 tokens per page):

MODEL	~ PAGES PER DOLLAR	PERFORMANCE ON MTEB EVAL	MAX INPUT
text-embedding-3-small	62,500	62.3%	8191
text-embedding-3-large	9,615	64.6%	8191
text-embedding-ada-002	12,500	61.0%	8191

Search & Rank Layer

This layer ensures that the retrieved text is accurate, relevant, and contextually appropriate. This layer consists of two main components: a search component that retrieves relevant documents using semantic similarity techniques, and a re-rank component that ranks the retrieved documents based on relevance, popularity, and freshness. The search and re-rank layer is crucial for tasks like question answering, summarization, and machine translation, as it helps in retrieving and ranking relevant information efficiently. It plays a key role in enhancing the performance of AI tasks and is an integral part of the project's RAG models.

Embeddings are generated using OpenAI and stored in ChromaDB. Embeddings of search query is generated using the same model and top 3 most matching results are obtained based on its cosine similarity.

```
# Set up the embedding function using the OpenAI embedding model

model = "text-embedding-ada-002"
embedding_function = OpenAIEmbeddingFunction(api_key=openai.api_key,
model_name=model)
```

```
# Initialise a collection in chroma and pass the embedding_function to it
so that it used OpenAI embeddings to embed the documents
insurance_collection = client.get_or_create_collection(name='RAG_on_Insur-
ance', embedding_function=embedding_function)
```

```
# Convert the page text and metadata from your dataframe to lists to be
able to pass it to chroma

documents_list = insurance_pdf_data["Page_Text"].tolist()
metadata_list = insurance_pdf_data['Metadata'].tolist()
```

```
# Add the documents and metadata to the collection alongwith generic inte-
ger IDs. You can also feed the metadata information as IDs by combining
the policy name and page no.

insurance_collection.add(
    documents= documents_list,
    ids = [str(i) for i in range(0, len(documents_list))],
    metadatas = metadata_list
)
```

Caching mechanism is also implemented to optimize search performance. Search query embedding and search result metadata is stored in Vector DB (ChromaDB in this case) whenever any search is performed.

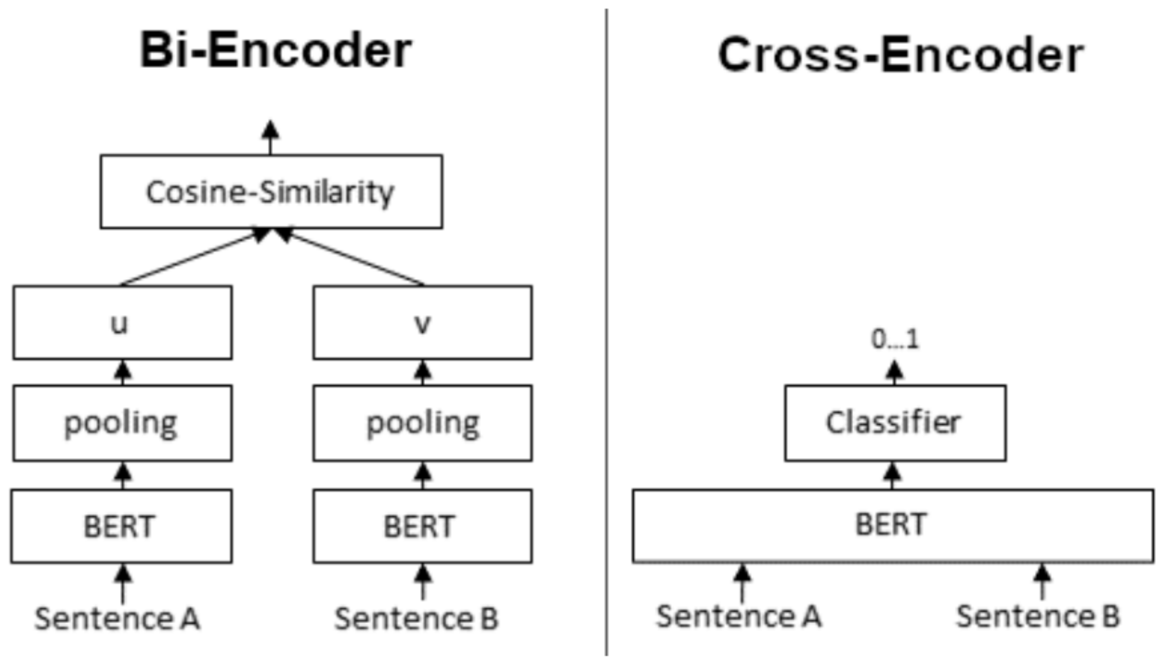
If new search query is found similar to already existing query in cache embeddings, depending on variance factor defined, then corresponding cached search results are used, rather than searching on whole document imbedding's.

```
# Implementing Cache in Semantic Search

# Set a threshold for cache search
threshold = 0.2

ids = []
documents = []
distances = []
metadatas = []
results_df = pd.DataFrame()


# If the distance is greater than the threshold, then return the results
from the main collection.

if cache_results['distances'][0] == [] or cache_results['distances'][0][0]
> threshold:
      # Query the collection against the user query and return the top 10
results
      results = insurance_collection.query(
      query_texts=query,
      n_results=10
      )

      # Store the query in cache_collection as document w.r.t to ChromaDB
so that it can be embedded and searched against later
      # Store retrieved text, ids, distances and metadatas in
cache_collection as metadatas, so that they can be fetched easily if a
query indeed matches to a query in cache
      Keys = []
      Values = []

      for key, val in results.items():
        if val is None:
          continue
        for i in range(10):
          Keys.append(str(key)+str(i))
          Values.append(str(val[0][i]))


      cache_collection.add(
          documents= [query],
          ids = [query],  # Or if you want to assign integers as IDs
0,1,2,.., then you can use "len(cache_results['documents'])" as will
```

Re-ranking:

Re-ranking the results obtained from semantic search can sometime significantly improve the relevance of the retrieved results. This is often done by passing the query paired with each of the retrieved responses into a cross-encoder to score the relevance of the response w.r.t. the query.

There are several cross-encoders available. We are using **cross-encoder/ms-marco-MiniLM-L-6-v2** from **sentence_transformers**



```
# Import the CrossEncoder library from sentence_transformers

from sentence_transformers import CrossEncoder, util
```

```
# Initialise the cross encoder model

cross_encoder = CrossEncoder('cross-encoder/ms-marco-MiniLM-L-6-v2')
```

```
# Input (query, response) pairs for each of the top 20 responses received
from the semantic search to the cross encoder
# Generate the cross_encoder scores for these pairs

cross_inputs = [[query, response] for response in results_df['Documents']]
cross_rerank_scores = cross_encoder.predict(cross_inputs)
```

```
# Store the rerank_scores in results_df

results_df['Reranked_scores'] = cross_rerank_scores
```

Semantic Search Result Code:

```
# Return the top 3 results after reranking

top_3_rerank = results_df.sort_values(by='Reranked_scores',
ascending=False)
top_3_rerank[:3]
```

```
top_3_RAG = top_3_rerank[["Documents", "Metadatas"]][:3]
```

Search Queries

1. **Que: What are the requirements for placing in force any Scheduled benefit that would have been subject to Proof of Good Health has the member remained continuously insured?**

```python
# Return the top 3 results after reranking

top_3_rerank = results_df.sort_values(by='Reranked_scores', ascending=False)
top_3_rerank[:3]
```

	Metadatas	Documents	Distances	IDs	Reranked_scores
4	{'Page_No.': 'Page 40'}	Section E - Reinstatement Article 1 - Reinstat...	0.269944	37	6.390748
1	{'Page_No.': 'Page 29'}	Insurance for which Proof of Good Health is re...	0.232257	26	4.750897
0	{'Page_No.': 'Page 31'}	Scheduled Benefit in force for the Member befo...	0.229051	28	4.389155

```python
top_3_RAG = top_3_rerank[["Documents", "Metadatas"]][:3]
```

```python
top_3_RAG
```

1 to 3 of 3 entries Filter

index	Documents	Metadatas
4	Section E - Reinstatement Article 1 - Reinstatement A Member's terminated insurance will be reinstated if: a. insurance ceased because of layoff or approved leave of absence; and b. the Member returns to Active Work for the Policyholder within six months of the date insurance ceased. The Member's reinstated insurance will be in force on the date of return to work. However, the Actively at Work and Period of Limited Activity provisions discussed in PART III, Section B, will apply. Also, Proof of Good Health will be required to place in force any Scheduled Benefit that would have been subject to Proof of Good Health had the Member remained continuously insured. Only the period of time during which a Member is actually insured will be included in determining the length of his or her continuous coverage under this Group Policy. For this purpose the period of time during which a reinstated Member's insurance was not in force: a. will not be considered an interruption of continuous coverage; and b. will not be used to satisfy any provision of this Group Policy which pertains to a period of continuous coverage. In addition, a longer reinstatement period may be allowed for an approved leave of absence taken in accordance with the provisions of the federal law regarding the Uniformed Services Employment and Reemployment Rights Act of 1994 (USERRA). Article 2 - Federal Required Family and Medical Leave Act (FMLA) A Member's terminated insurance may be reinstated in accordance with the provisions of the Federal Family and Medical Leave Act (FMLA), subject to the Actively at Work and Period of Limited Activity provision discussed in PART III, Section B. Article 3 - Reinstatement of Coverage for a Member or Dependent When Coverage Ends due to Living Outside of the United States This policy has been updated effective January 1, 2014 PART III - INDIVIDUAL REQUIREMENTS AND RIGHTS GC 6010 Section E - Reinstatement, Page 1	{'Page_No.': 'Page 40'}
1	Insurance for which Proof of Good Health is required (see e. below) will be in force on the later of: (1) the date insurance would have been effective if Proof of Good Health had not been required; or (2) the date Proof of Good Health is approved by The Principal. c. Effective Date for Initial Noncontributory Insurance When Proof of Good Health is not Required Unless Proof of Good Health is required (see b. above and e. below), insurance for which the Member contributes no part of premium will be in force on the date the Member is eligible. d. Effective Date for Initial Contributory Insurance When Proof of Good Health is not Required If a Member is to contribute a part of premium, insurance must be requested in a form provided by The Principal. Unless Proof of Good Health is required (see b. above and e. below), the requested insurance will be in force on: (1) the date the Member is eligible, if the request is made on or before that date; or (2) the date of the Member's request, if the request is made within 31 days after the date the Member is eligible. If the request is made more than 31 days after the date the Member is eligible, Proof of Good Health will be required before insurance can be in force (see b. above and e. below). e. Proof of Good Health Requirements The type and form of required Proof of Good Health will be determined by The Principal. A Member must submit Proof of Good Health: (1) If insurance for which a Member contributes a part of premium is requested more than 31 days after the date the Member is eligible including any insurance the Member refuses and later requests. (2) If insurance is requested under this Group Policy by a Member that was eligible under the Prior Policy, but elected to waive coverage under the Prior Policy. (3) If a Member has failed to provide required Proof of Good Health or has been refused insurance under this Group Policy at any prior time. (4) If a Member elects to terminate insurance and, more than 31 days later, requests to be insured again. (5) If, on the date a Member becomes eligible, fewer than five Members are insured. This policy has been updated effective January 1, 2014 PART III - INDIVIDUAL REQUIREMENTS AND RIGHTS GC 6007 Section B - Effective Dates, Page 2	{'Page_No.': 'Page 29'}
0	Scheduled Benefit in force for the Member before the change will continue to apply to the Member until the day of return to Active Work. When the Member returns to Active Work, the Scheduled Benefit increase will then be in force for the Member. Any decrease in Scheduled Benefit amounts due to a change by amendment to this Group Policy will be effective on the date of change, whether or not the Member is Actively at Work. (2) A change in the Member's Scheduled Benefit amount because of a change in the Schedule of Insurance (as described in PART IV, Section A) by amendment to this Group Policy for which Proof of Good Health is required (see e. above) will be effective on the later of: - the date the change would have been effective if Proof of Good Health had not been required; or - the date Proof of Good Health is approved by The Principal. h. Effective Date for Benefit Changes Due to Changes Requested by the Member (1) A change in a Member's Scheduled Benefit amount because of a request by the Member for which Proof of Good Health is not required (see e. above) will normally be effective on the date of the request. However, if the Member is not Actively at Work on the date a Scheduled Benefit change would otherwise be effective, the Scheduled Benefit change will not be in force until the date the Member returns to Active Work. Any decrease in Scheduled Benefit amounts due to a request by the Member will be effective on the date of the change, whether or not the Member is Actively at Work. (2) A change in the Member's Scheduled Benefit amount because of a request by the Member for which Proof of Good Health is required (see e. above) will be effective on the later of: - the date the change would have been effective if Proof of Good Health had not been required; or - the date Proof of Good Health is approved by The Principal. i. Effective Date for Benefit Changes Due to Change in the Member's Family Status A Member may request an increase in Scheduled Benefit amount, a decrease in Scheduled Benefits, or the addition of Scheduled Benefits for which he or she was not previously insured if a change in the Member's family status as described below has occurred, provided a request for such increase, decrease, or addition is made in Writing within 31 days after the date of the change in family status: This policy has been updated effective January 1, 2014 PART III - INDIVIDUAL REQUIREMENTS AND RIGHTS GC 6007 Section B - Effective Dates, Page 4	{'Page_No.': 'Page 31'}

Answer is clearly found in first result which is mentioned on page no 40 of the document.

2. **Que: How is the peroid of time during which a reinstated Member's insurance was not in force treated for the purpose of determining the length of continuous coverage under the Group Policy?**

```python
# Return the top 3 results after reranking
top_3_rerank = results_df.sort_values(by='Reranked_scores', ascending=False)
top_3_rerank[:3]
```

	Metadatas	Documents	Distances	IDs	Reranked_scores
0	{'Page_No.': 'Page 40'}	Section E - Reinstatement Article 1 - Reinstat...	0.214788	37	4.639090
1	{'Page_No.': 'Page 41'}	I f coverage for a Member or Dependent termina...	0.258539	38	-1.227318
5	{'Page_No.': 'Page 33'}	a . In no event will Dependent Life Insurance ...	0.326898	30	-1.711988

```python
top_3_RAG = top_3_rerank[["Documents", "Metadatas"]][:3]
```

```python
top_3_RAG
```

1 to 3 of 3 entries Filter

index	Documents	Metadatas
4	Section E - Reinstatement Article 1 - Reinstatement A Member's terminated insurance will be reinstated if: a. insurance ceased because of layoff or approved leave of absence; and b. the Member returns to Active Work for the Policyholder within six months of the date insurance ceased. The Member's reinstated insurance will be in force on the date of return to work. However, the Actively at Work and Period of Limited Activity provisions discussed in PART III, Section B, will apply. Also, Proof of Good Health will be required to place in force any Scheduled Benefit that would have been subject to Proof of Good Health had the Member remained continuously insured. Only the period of time during which a Member is actually insured will be included in determining the length of his or her continuous coverage under this Group Policy. For this purpose the period of time during which a reinstated Member's insurance was not in force: a. will not be considered an interruption of continuous coverage; and b. will not be used to satisfy any provision of this Group Policy which pertains to a period of continuous coverage. In addition, a longer reinstatement period may be allowed for an approved leave of absence taken in accordance with the provisions of the federal law regarding the Uniformed Services Employment and Reemployment Rights Act of 1994 (USERRA). Article 2 - Federal Required Family and Medical Leave Act (FMLA) A Member's terminated insurance may be reinstated in accordance with the provisions of the Federal Family and Medical Leave Act (FMLA), subject to the Actively at Work and Period of Limited Activity provision discussed in PART III, Section B. Article 3 - Reinstatement of Coverage for a Member or Dependent When Coverage Ends due to Living Outside of the United States This policy has been updated effective January 1, 2014 PART III - INDIVIDUAL REQUIREMENTS AND RIGHTS GC 6010 Section E - Reinstatement, Page 1	{'Page_No.': 'Page 40'}
1	Insurance for which Proof of Good Health is required (see e. below) will be in force on the later of: (1) the date insurance would have been effective if Proof of Good Health had not been required; or (2) the date Proof of Good Health is approved by The Principal. c. Effective Date for Initial Noncontributory Insurance When Proof of Good Health is not Required Unless Proof of Good Health is required (see b. above and e. below), insurance for which the Member contributes no part of premium will be in force on the date the Member is eligible. d. Effective Date for Initial Contributory Insurance When Proof of Good Health is not Required If a Member is to contribute a part of premium, insurance must be requested in a form provided by The Principal. Unless Proof of Good Health is required (see b. above and e. below), the requested insurance will be in force on: (1) the date the Member is eligible, if the request is made on or before that date; or (2) the date of the Member's request, if the request is made within 31 days after the date the Member is eligible. If the request is made more than 31 days after the date the Member is eligible, Proof of Good Health will be required before insurance can be in force (see b. above and e. below). e. Proof of Good Health Requirements The type and form of required Proof of Good Health will be determined by The Principal. A Member must submit Proof of Good Health: (1) If insurance for which a Member contributes a part of premium is requested more than 31 days after the date the Member is eligible including any insurance the Member refuses and later requests. (2) If insurance is requested under this Group Policy by a Member that was eligible under the Prior Policy, but elected to waive coverage under the Prior Policy. (3) If a Member has failed to provide required Proof of Good Health or has been refused insurance under this Group Policy at any prior time. (4) If a Member elects to terminate insurance and, more than 31 days later, requests to be insured again. (5) If, on the date a Member becomes eligible, fewer than five Members are insured. This policy has been updated effective January 1, 2014 PART III - INDIVIDUAL REQUIREMENTS AND RIGHTS GC 6007 Section B - Effective Dates, Page 2	{'Page_No.': 'Page 29'}
0	Scheduled Benefit in force for the Member before the change will continue to apply to the Member until the day of return to Active Work. When the Member returns to Active Work, the Scheduled Benefit increase will then be in force for the Member. Any decrease in Scheduled Benefit amounts due to a change by amendment to this Group Policy will be effective on the date of change, whether or not the Member is Actively at Work. (2) A change in the Member's Scheduled Benefit amount because of a change in the Schedule of Insurance (as described in PART IV, Section A) by amendment to this Group Policy for which Proof of Good Health is required (see e. above) will be effective on the later of: - the date the change would have been effective if Proof of Good Health had not been required; or - the date Proof of Good Health is approved by The Principal. h. Effective Date for Benefit Changes Due to Changes Requested by the Member (1) A change in a Member's Scheduled Benefit amount because of a request by the Member for which Proof of Good Health is not required (see e. above) will normally be effective on the date of the request. However, if the Member is not Actively at Work on the date a Scheduled Benefit change would otherwise be effective, the Scheduled Benefit change will not be in force until the date the Member returns to Active Work. Any decrease in Scheduled Benefit amounts due to a request by the Member will be effective on the date of the change, whether or not the Member is Actively at Work. (2) A change in the Member's Scheduled Benefit amount because of a request by the Member for which Proof of Good Health is required (see e. above) will be effective on the later of: - the date the change would have been effective if Proof of Good Health had not been required; or - the date Proof of Good Health is approved by The Principal. i. Effective Date for Benefit Changes Due to Change in the Member's Family Status A Member may request an increase in Scheduled Benefit amount, a decrease in Scheduled Benefits, or the addition of Scheduled Benefits for which he or she was not previously insured if a change in the Member's family status as described below has occurred, provided a request for such increase, decrease, or addition is made in Writing within 31 days after the date of the change in family status: This policy has been updated effective January 1, 2014 PART III - INDIVIDUAL REQUIREMENTS AND RIGHTS GC 6007 Section B - Effective Dates, Page 4	{'Page_No.': 'Page 31'}

Answer is clearly found in first result which is mentioned on page no 40 of the document.

3. **What provisions may allow for a longer reinstatement period for an approved leave of absence taken in accordance with the Uniformed Services Employment and Reemployment Rights Act of 1994 (USERRA)?**

```python
# Return the top 3 results after reranking
top_3_rerank = results_df.sort_values(by='Reranked_scores', ascending=False)
top_3_rerank[:3]
```

	Metadatas	Documents	Distances	IDs	Reranked_scores
0	{'Page_No.': 'Page 40'}	Section E - Reinstatement Article 1 - Reinstat...	0.263336	37	5.194175
3	{'Page_No.': 'Page 38'}	Section D - Continuation Article 1 - Member Li...	0.417215	35	-0.008611
1	{'Page_No.': 'Page 41'}	I f coverage for a Member or Dependent termina...	0.353420	38	-3.647968

```
top_3_RAG = top_3_rerank[["Documents", "Metadatas"]][:3]
```

```
[97] top_3_RAG
```

1 to 3 of 3 entries Filter

index	Documents	Metadatas
4	Section E - Reinstatement Article 1 - Reinstatement A Member's terminated insurance will be reinstated if: a. insurance ceased because of layoff or approved leave of absence; and b. the Member returns to Active Work for the Policyholder within six months of the date insurance ceased. The Member's reinstated insurance will be in force on the date of return to work. However, the Actively at Work and Period of Limited Activity provisions discussed in PART III, Section B, will apply. Also, Proof of Good Health will be required to place in force any Scheduled Benefit that would have been subject to Proof of Good Health had the Member remained continuously insured. Only the period of time during which a Member is actually insured will be included in determining the length of his or her continuous coverage under this Group Policy. For this purpose the period of time during which a reinstated Member's insurance was not in force: a. will not be considered an interruption of continuous coverage; and b. will not be used to satisfy any provision of this Group Policy which pertains to a period of continuous coverage. In addition, a longer reinstatement period may be allowed for an approved leave of absence taken in accordance with the provisions of the federal law regarding the Uniformed Services Employment and Reemployment Rights Act of 1994 (USERRA). Article 2 - Federal Required Family and Medical Leave Act (FMLA) A Member's terminated insurance may be reinstated in accordance with the provisions of the Federal Family and Medical Leave Act (FMLA), subject to the Actively at Work and Period of Limited Activity provision discussed in PART III, Section B. Article 3 - Reinstatement of Coverage for a Member or Dependent When Coverage Ends due to Living Outside of the United States This policy has been updated effective January 1, 2014 PART III - INDIVIDUAL REQUIREMENTS AND RIGHTS GC 6010 Section E - Reinstatement, Page 1	{'Page_No.': 'Page 40'}
1	Insurance for which Proof of Good Health is required (see e. below) will be in force on the later of: (1) the date insurance would have been effective if Proof of Good Health had not been required; or (2) the date Proof of Good Health is approved by The Principal. c. Effective Date for Initial Noncontributory Insurance When Proof of Good Health is not Required Unless Proof of Good Health is required (see b. above and e. below), insurance for which the Member contributes no part of premium will be in force on the date the Member is eligible. d. Effective Date for Initial Contributory Insurance When Proof of Good Health is not Required If a Member is to contribute a part of premium, insurance must be requested in a form provided by The Principal. Unless Proof of Good Health is required (see b. above and e. below), the requested insurance will be in force on: (1) the date the Member is eligible, if the request is made on or before that date; or (2) the date of the Member's request, if the request is made within 31 days after the date the Member is eligible. If the request is made more than 31 days after the date the Member is eligible, Proof of Good Health will be required before insurance can be in force (see b. above and e. below). e. Proof of Good Health Requirements The type and form of required Proof of Good Health will be determined by The Principal. A Member must submit Proof of Good Health: (1) If insurance for which a Member contributes a part of premium is requested more than 31 days after the date the Member is eligible including any insurance the Member refuses and later requests. (2) If insurance is requested under this Group Policy by a Member that was eligible under the Prior Policy, but elected to waive coverage under the Prior Policy. (3) If a Member has failed to provide required Proof of Good Health or has been refused insurance under this Group Policy at any prior time. (4) If a Member elects to terminate insurance and, more than 31 days later, requests to be insured again. (5) If, on the date a Member becomes eligible, fewer than five Members are insured. This policy has been updated effective January 1, 2014 PART III - INDIVIDUAL REQUIREMENTS AND RIGHTS GC 6007 Section B - Effective Dates, Page 2	{'Page_No.': 'Page 29'}
0	Scheduled Benefit in force for the Member before the change will continue to apply to the Member until the day of return to Active Work. When the Member returns to Active Work, the Scheduled Benefit increase will then be in force for the Member. Any decrease in Scheduled Benefit amounts due to a change by amendment to this Group Policy will be effective on the date of change, whether or not the Member is Actively at Work. (2) A change in the Member's Scheduled Benefit amount because of a change in the Schedule of Insurance (as described in PART IV, Section A) by amendment to this Group Policy for which Proof of Good Health is required (see e. above) will be effective on the later of: - the date the change would have been effective if Proof of Good Health had not been required; or - the date Proof of Good Health is approved by The Principal. h. Effective Date for Benefit Changes Due to Changes Requested by the Member (1) A change in a Member's Scheduled Benefit amount because of a request by the Member for which Proof of Good Health is not required (see e. above) will normally be effective on the date of the request. However, if the Member is not Actively at Work on the date a Scheduled Benefit change would otherwise be effective, the Scheduled Benefit change will not be in force until the date the Member returns to Active Work. Any decrease in Scheduled Benefit amounts due to a request by the Member will be effective on the date of the change, whether or not the Member is Actively at Work. (2) A change in the Member's Scheduled Benefit amount because of a request by the Member for which Proof of Good Health is required (see e. above) will be effective on the later of: - the date the change would have been effective if Proof of Good Health had not been required; or - the date Proof of Good Health is approved by The Principal. i. Effective Date for Benefit Changes Due to Change in the Member's Family Status A Member may request an increase in Scheduled Benefit amount, a decrease in Scheduled Benefits, or the addition of Scheduled Benefits for which he or she was not previously insured if a change in the Member's family status as described below has occurred, provided a request for such increase, decrease, or addition is made in Writing within 31 days after the date of the change in family status: This policy has been updated effective January 1, 2014 PART III - INDIVIDUAL REQUIREMENTS AND RIGHTS GC 6007 Section B - Effective Dates, Page 4	{'Page_No.': 'Page 31'}

Answer is clearly found in first result which is mentioned on page no 40 of the document.

Generation Layer

This is the final stage of the RAG model, featuring a large language model trained on extensive text and code datasets. It enables the model to produce new text in response to user queries. This layer synthesizes retrieved information, shaping it into coherent and contextually relevant responses. Crucial for tasks like question answering, summarization, machine translation, and generative search, the generation layer excels in providing context and natural language capabilities for such searches.

Find below prompt used. This also mentioned **Few Shots Prompts** for better results.

```python
# Define the function to generate the response. Provide a comprehensive
prompt that passes the user query and the top 3 results to the model

def generate_response(query, results_df):
    """
    Generate a response using GPT-3.5's ChatCompletion based on the user
query and retrieved information.
    """
    messages = [
                {"role": "system", "content":  "You are a helpful assis-
tant in the insurance domain who can effectively answer user queries about
insurance policies and documents."},
                {"role": "user", "content": f"""You are a helpful assis-
tant in the insurance domain who can effectively answer user queries about
insurance policies and documents.
                                                You have a question asked
by the user in '{query}' and you have some search results from a corpus of
insurance documents in the dataframe '{top_3_RAG}'. These search results
are essentially one page of an insurance document that may be relevant to
the user query.

                                                The column 'documents' in-
side this dataframe contains the actual text from the policy document and
the column 'metadata' contains source page. The text inside the document
may also contain tables in the format of a list of lists where each of the
nested lists indicates a row.

                                                Use the documents in
'{top_3_RAG}' to answer the query '{query}'. Frame an informative answer
and also, use the dataframe to return page numbers as citations.

                                                Examples:
                                                ####
                                                User Message: Under what
four conditions may a member's insurance be continued if Active work ends
due to layoffs or approved leave of absence?
                                                Assistant: Based on the
provided documents, the four conditions under which a member's insurance
may be continued if Active work ends due to layoffs or approved leave of
absence are:
                                                1. Continuation coverage
for a specific period if the member is laid off.
                                                2. Continuation coverage
for a specific period if the member is on an approved leave of absence.
```

Query Outputs:

Query 1: What are the requirements for placing in force any Scheduled benefit that would have been subject to Proof of Good Health has the member remained continuously insured?

```
# Print the response
print("\n".join(response))
```

```
Based on the provided documents, the requirements for placing in force any Scheduled benefit that would have been subject to Proof of Good Health if the member remained continuously insu

1. The Scheduled Benefit must have been in force for the Member before the Proof of Good Health requirement.
2. The Member must have remained continuously insured without any break in coverage.
3. The Scheduled Benefit must meet any other policy-specific criteria outlined in the insurance documents.

Citations:
- Scheduled Benefit in force for the Member before Proof of Good Health: Page 31
- Additional policy-specific criteria: Page 29, Page 40
```

```
Based on the provided documents, the requirements for placing in force any
Scheduled benefit that would have been subject to Proof of Good Health if
the member remained continuously insured are as follows:

1. The Scheduled Benefit must have been in force for the Member before the
Proof of Good Health requirement.
2. The Member must have remained continuously insured without any break in
coverage.
3. The Scheduled Benefit must meet any other policy-specific criteria out-
lined in the insurance documents.

Citations:
- Scheduled Benefit in force for the Member before Proof of Good Health:
Page 31
- Additional policy-specific criteria: Page 29, Page 40
```

Query 2: How is the period of time during which a reinstated Member's insurance was not in force treated for the purpose of determining the length of continuous coverage under the Group Policy?

```
# Print the response
print("\n".join(response))
```

```
Based on the documents provided, the period of time during which a reinstated Member's insurance was not in force is typically treated as follows for the purpose of determining the lengt

When a Member's insurance is reinstated, the period of time during which their insurance was not in force may not be counted towards the calculation of continuous coverage.

This means that the time when the insurance was not active might not be included in determining the length of continuous coverage under the Group Policy. However, it is essential to refe

Citations:
- Section E - Reinstatement Article 1: Page 40
- 'Page 41'
```

Based on the documents provided, the period of time during which a reinstated Member's insurance was not in force is typically treated as follows for the purpose of determining the length of continuous coverage under the Group Policy:

When a Member's insurance is reinstated, the period of time during which their insurance was not in force may not be counted towards the calculation of continuous coverage.

This means that the time when the insurance was not active might not be included in determining the length of continuous coverage under the Group Policy. However, it is essential to refer to the specific details outlined in the insurance policy document to understand the exact treatment of this period.

Citations:
- Section E - Reinstatement Article 1: Page 40
- 'Page 41'

Query 3: What provisions may allow for a longer reinstatement period for an approved leave of absence taken in accordance with the Uniformed Services Employment and Reemployment Rights Act of 1994 (USERRA)?

```
# Print the response
print("\n".join(response))
```

```
Based on the provided documents, the provisions that may allow for a longer reinstatement period for an approved leave of absence taken in accordance with the Uniformed Services Employme

Citations:
- Section E - Reinstatement, Page 40
```

Based on the provided documents, the provisions that may allow for a longer reinstatement period for an approved leave of absence taken in accordance with the Uniformed Services Employment and Reemployment Rights Act of 1994 (USERRA) can be found in the Section E - Reinstatement. The detailed explanation regarding the reinstatement period extension, any specific conditions or criteria for eligibility, and the process for reinstatement after an approved leave of absence under USERRA would be specified in this section.

Citations:
- Section E - Reinstatement, Page 40