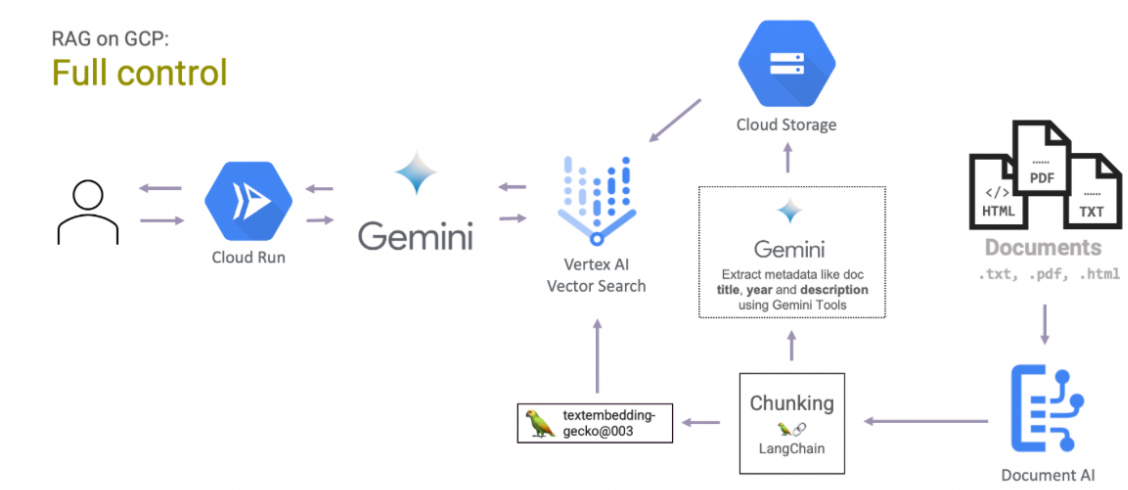
Problema 1:

First of all, we can use the **Google Cloud Storage** to store the documents. We can use **Flask** to be able to upload and delete the documents stored in the Cloud Storage and get the input from the user and provide an output. We can also create a pipeline that converts the pdf into a plain text and cleans the unimportant parts of the file depending on the task. After obtaining the plain text, we need to separate the text into chunks using **semantic segmentation** with the help of the libraries such as **NLTK**. After dividing the sentences we need to embed them into vectors, we need to use a **pre-trained embedding model** like **BERT** and store them into a vector database. By using the same embedding model, we need to **embed the input** retrieved from the user and do a **cosine-similarity** search to gather the most similar chunks in our vector storage. After getting the most similar chunks, we need to plug the chunks into a **generative model** such as gpt or gemini for it to generate an answer **based on the documents** avaible in the **Cloud Storage**.



Problema 2:

First of all, we need to use **speech-to-text transformation** and store the data, we can use **Google Cloud speech to text**. Then we can fine-tune an LLM to do a **topic extraction**, based on the call centers topic, from each call and categorize the calls by topics. We can use **topic clustering** by defining a similarity score to identify each topic, such as if topic 1 and topic 16 has a similarity score bigger then %n then we can combine these topics under one topic to avoid confusions. As we will examine a call center data we can not just use a **sentiment analysis** on the whole conversation as each call is made because of a problem. We can make a sentiment analysis on the **last few exchanges** made in a call to see how the call has ended without using a survey. This way we can extract the sentiments of the closing argument of the caller, and conclude if they are satisfied or dissatisfied by the call. For the sentiment analysis, we can fine tune a pre-trained sentiment analysis model to make sure it understands how to analyse a call. After performing all the operations we can provide graphics regarding the call center.