<u>*LLM Project (Deadline – 30th Nov, 2024)*</u>

You have been tasked with architecting a system to build and manage a question-answering tool that can help users answer questions about legal contracts. The intended audience for this tool is a user at a company who uses the question-answering tool to help facilitate document review.

You do not need to implement every component yourself. Feel free to use a third-party LLM via local inference or an API to a cloud service as you see fit if you would like. Focus on the general ML system design thinking and implementation abilities.

Explain your rationality of choosing any very specific LLM.

Please implement a simple question-answering system that satisfies the following requirements:

1. Reads a provided PDF input file. It's ok to assume the PDF has readable text in it - you do not need to implement OCR functionality.
2. Allows a user to answer a question about the document in natural language form
3. Cites the source within the document when giving an answer. Do not simply assert a claim.

Implementation

- Please implement your project as a command-line interface written in Python. You can use other packages and cloud services as you see fit, but be prepared to discuss the design decisions you made in selecting the external dependencies.

Hint: *Follow the link below and design a similar system but don't copy every line. Also implement with the help of Classes and Methods. And Question answer system should be command line prompts*

https://github.com/aniketpotabatti/Gemini-PDF-Question-Answering-System

Interface

python3 question-answering.py /path/to/document "question goes here"

Output

answer in natural language form

Requirements

1. Ensure the README.md file has:
    1. Sufficient information for an evaluating engineer to run the project
    2. Answers to the following questions:
        i. Briefly describe the architecture of your approach
        ii. What are the major pitfalls of your design?
        iii. What are some safeguards you would implement if you were to develop a commercial product?