Name-Vidwesh Choudhary

Steps which I followed to do this project are:

1.OCR Configuration: Tesseract OCR can be used to extract Gujarati text from PDFs because they sometimes contain scanned images rather than selected text.  
2.Convert PDF to Images: Because pdf2image ensures precise text extraction from scanned documents by converting each page into an image for OCR processing.  
3.Text extraction to JSON: To make searching and retrieving the extracted text easier, store it in an organized fashion.  
4.Why the Flask API was created: Flask offers a small server that can effectively manage Q&A queries.  
5.Find Answers: Look for user-provided queries in the JSON and deliver pertinent data for a simple Q&A system.