**Intelligent Form Agent**

**Overview**

The Intelligent Form Agent is a web-based application built using Streamlit that streamlines the process of analysing documents. Traditionally, reviewing multiple forms, resumes, or reports is time-consuming and prone to human error. This agent automates the workflow by allowing users to upload documents in various formats including PDFs, DOCX files, and images and efficiently extract meaningful content from them. Once the documents are uploaded, the agent performs text extraction, generates per-document summaries, creates a holistic summary across all uploaded documents, and enables interactive querying via a Q&A interface.

The key objective of the Intelligent Form Agent is to make document analysis faster, more accurate, and user-friendly. Users can not only view concise summaries of each file but also ask specific questions about the content. The system intelligently consolidates information from multiple documents to provide accurate answers, making it suitable for HR teams, data analysts, and researchers who need to process and analyse large volumes of structured or unstructured text data.

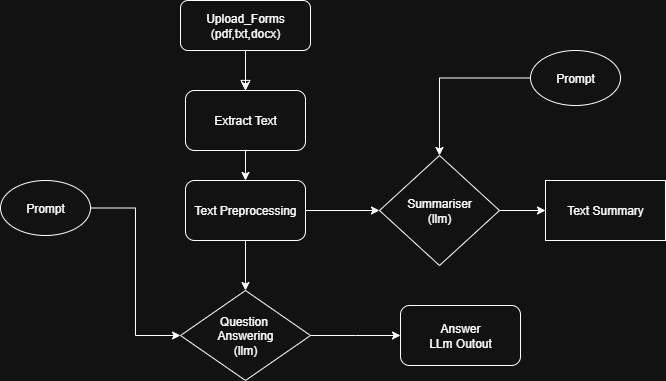
**Features**

* **Text Extraction:** Supports PDF, DOCX, JPG, PNG files.
* **Per-Document Summaries:** Generates concise summaries for each uploaded document.
* **Holistic Summary:** Combines multiple documents into a single summary for better insight.
* **Interactive Q&A:** Users can ask questions about the content in a sidebar chat interface.
* **Session Persistence:** Uploaded files, summaries, and holistic summaries are cached in session state.

**Architecture**

* **Frontend:** Streamlit for file upload, summaries display, and Q&A interface.
* **Backend:**
  + extract\_multiple\_forms – Handles text extraction from documents.
  + generate\_text\_summary – Generates per-document summaries using Groq API.
  + generate\_holistic\_summary – Combines all extracted text for a single holistic summary.
  + llm\_reply – Handles question-answering over extracted content.
* **State Management:** Streamlit session state to maintain uploaded files, extracted text, and summaries.

**Flowchart**

****

**Usage**

1. Upload one or more documents (PDF, DOCX, images).
2. View per-document summaries under each file.
3. If multiple files are uploaded, view the holistic summary.
4. Ask questions in the sidebar; answers are generated using combined context if multiple files are uploaded.

**Example Queries**

1. “List key skills of each individual.”
2. “Who has done the most projects?”
3. “Summarize education and work experience.”

