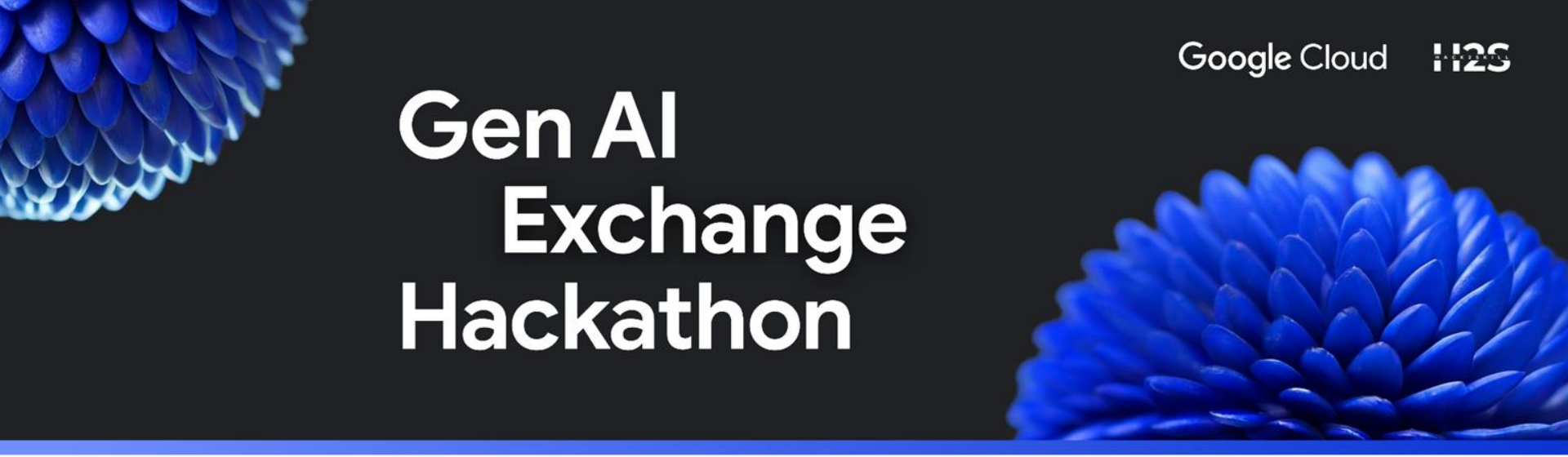


Team Name : Coding Wizards

Team Leader Name : Himanshi

Problem Statement : Generative AI for Demystifying Legal Documents

Brief about the prototype:

Our AI-Powered Legal Tool
This tool is designed as a **user-friendly application** that simplifies **complex legal documents** and makes them understandable for everyday users. It serves as a **practical, interactive tool** capable of summarizing legal text, answering clause-specific questions, and providing **voice-enabled guidance**. By combining **AI-powered summarization, interactive Q&A, multilingual support, and privacy protection**, the tool empowers users—**citizens, small business owners, non-English speakers, and accessibility-dependent individuals**—to make informed decisions without needing a legal background.

Core Functionalities of the Tool:

1) **Document Processing:** Accepts **PDF, DOCX, and image files** for flexible input.
2) **Text Extraction & Analysis:** Extracts all text accurately using **pdfplumber, docx2txt, and Tesseract OCR**, including from scanned images.
3) **Intelligent Summarization:** Converts complex legal clauses into **plain-language summaries** with **flan-t5-large** and **legal-summarizer-bart**, making content easy to understand.
4) **Voice-Friendly Refinement: Gemini LLM** refines summaries into **conversational text**, suitable for reading or audio generation.
5) **Interactive Q&A:** Users can ask **clause-specific questions**, receiving **context-aware answers** via **all-MiniLM-L6-v2 embeddings** and **FLAN-T5-small**.
6) **Speech Integration:** Generates **high-quality voice output** using **Coqui TTS** for accessibility.
7) **Privacy Protection:** Automatically masks **sensitive information** (names, IDs, addresses) via **Presidio** to ensure secure use.
8) **This tool** provides a **complete end-to-end solution**, from uploading legal documents to receiving **readable summaries, interactive answers, and audio outputs**, making legal guidance **accessible, secure, and actionable**.

Opportunity Analysis

❑ How Our Tool is Different From Existing Solutions

Many current solutions for legal document understanding have **significant limitations**:

1) **ChatGPT:** Provides general AI responses but **lacks specialization in legal documents** and does not support **voice interaction**.
2) **Legal AI SaaS Apps:** Offer document analysis but often **miss comprehensive voice support** and **PII masking**, leaving privacy concerns unaddressed.
3) **Traditional Tools:** Summarize or extract content but do not offer an **end-to-end solution** combining **summarization, clause-level Q&A, and speech output**.

❑ How Our Tool Solves the Problem:

1) **Accessibility:** Converts dense legal documents into **plain-language summaries**, making content understandable for non-experts.
2) **Voice Interaction:** Allows **hands-free review** of legal content, generating **audio outputs** from summaries and Q&A responses.
3) **Privacy Protection: Sensitive information is masked** (names, IDs, addresses) before AI processing, ensuring **user confidentiality**.
4) **Comprehensive Understanding:** Integrates **summarization, interactive Q&A, and speech**, providing a **holistic view** of legal documents in one tool.

❑ Unique Selling Proposition (USP): Three-Pillar Differentiation

1) **Multilingual Support:** Leverages **Gemini's multilingual capabilities** for input and output in multiple languages.
2) **Complete Pipeline:** Seamlessly integrates **Upload → Simplify → Ask → Listen**, offering an **end-to-end experience**.
3) **Privacy-First Design:** Uses **local models** for most processing and **controlled Gemini usage**, ensuring **maximum data security**.

List of features offered by the solution

1) **Document Upload & Compatibility:** Users can upload **PDF, DOCX, and image-based legal documents**, making the system flexible for various formats.

2) **Privacy & Security:** Automatically masks sensitive data such as **names, IDs, addresses, and contact details** using **Presidio**, ensuring **confidentiality and secure processing**.

3) **Text Extraction:** Accurately extracts content from all document types using **pdfplumber, docx2txt, and Tesseract OCR**, ensuring no important information is missed.

4) **Advanced Summarization:** Converts complex legal language into **clear, concise, and readable summaries** using **flan-t5-large** and **legal-summarizer-bart**, making documents accessible to non-experts.

5) **Clause-Level Question & Answer:** Allows users to **ask specific questions about clauses**; responses are generated using **all-MiniLM-L6-v2 embeddings** and **FLAN-T5-small**, delivering **context-aware, precise answers**.

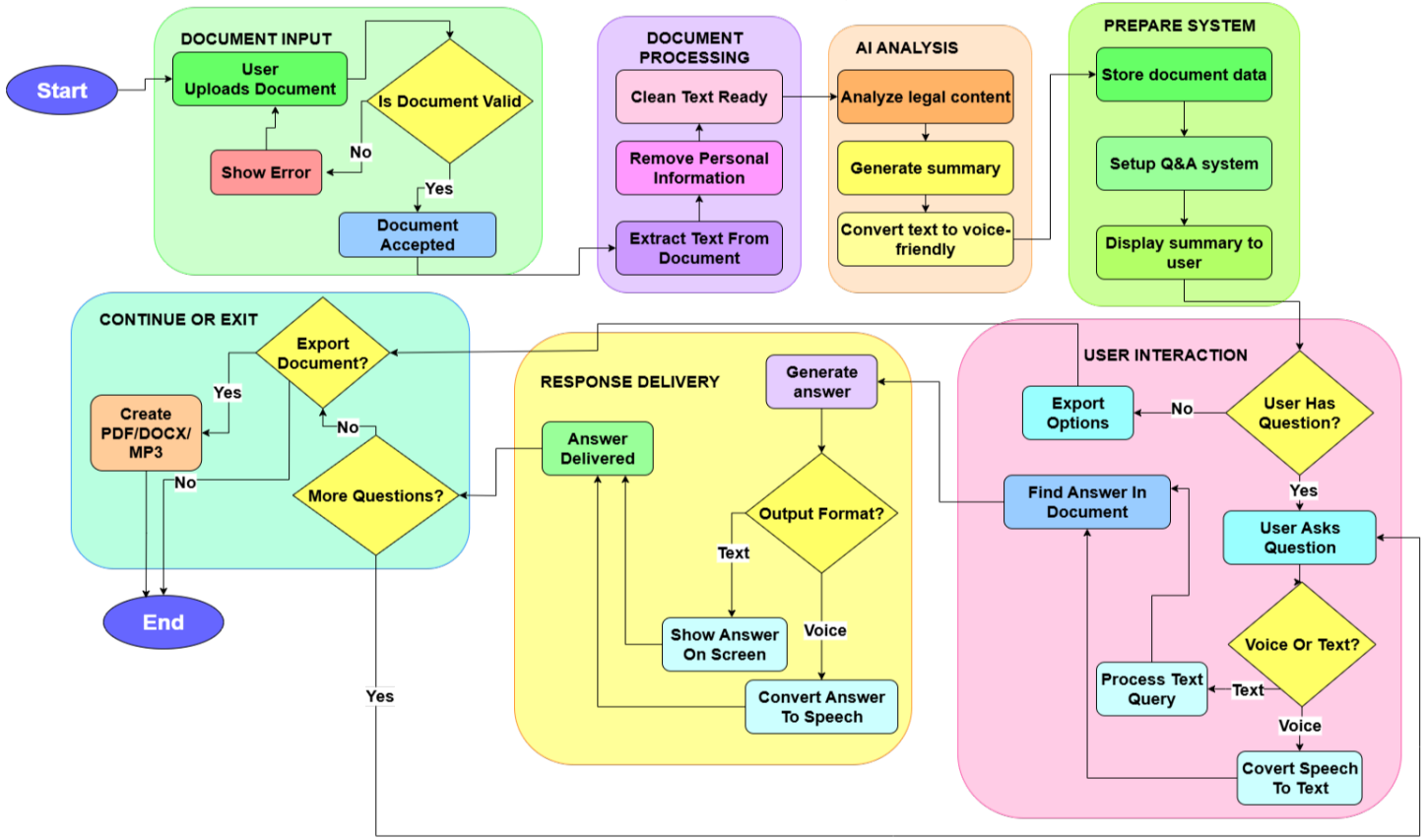
6) **Text Refinement for Accessibility:** Summaries are refined with **Gemini LLM** to produce **spoken-friendly, easy-to-read text** suitable for both reading and audio output.

7) **Voice Support:** Generates **high-quality audio responses** using **Coqui TTS**, enabling users to listen to legal explanations for better understanding.

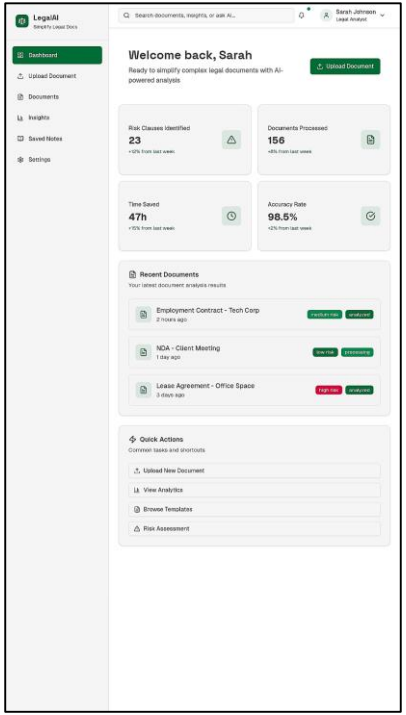
8) **Multilingual Support:** Accepts input and produces output in **multiple languages**, expanding accessibility to non-English speaking users.

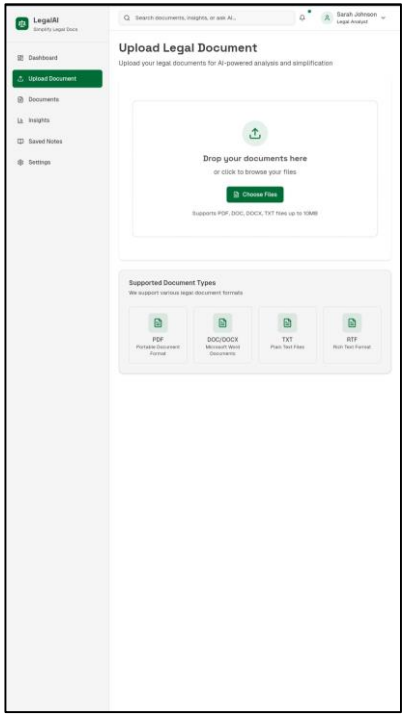
9) **Export Options:** Users can **download summaries, Q&A, and audio files** as PDF, DOCX, or MP3 for **offline reference**.

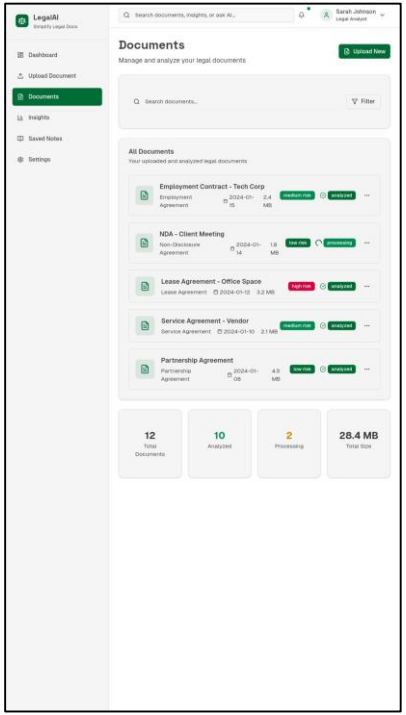
Process flow diagram

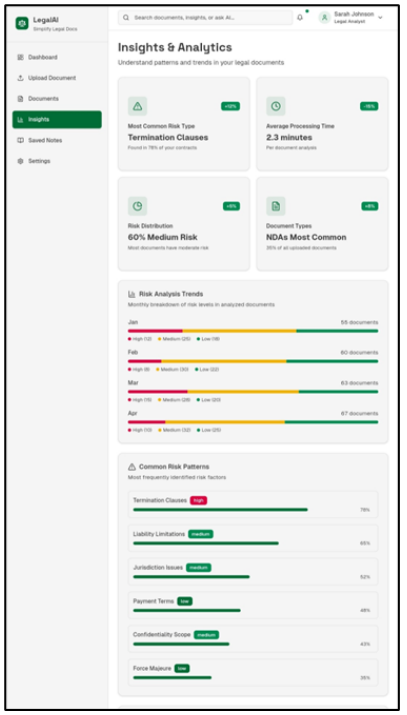


Wireframes of the proposed solution

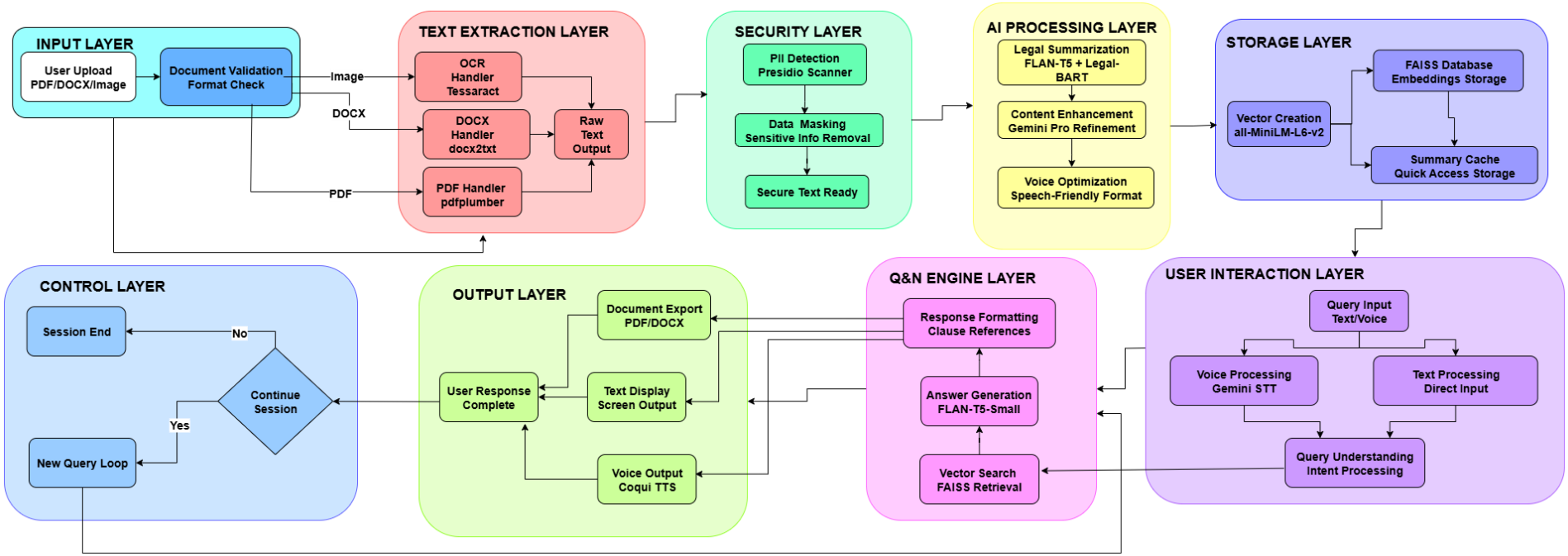








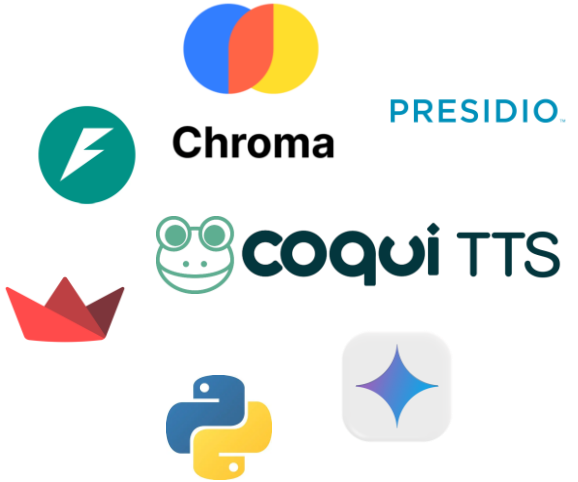
Architecture diagram of the proposed solution



Technologies to be used in the solution:

Our solution integrates a **balanced tech stack** that ensures efficiency, accessibility, and security

- ❑ **Programming Language & Backend:** Python with lightweight frameworks (FastAPI / Flask) to power APIs, manage uploads, and connect AI modules.
- ❑ **Frontend:** Streamlit for a simple, interactive web interface enabling **document upload, multilingual options, Q&A, and audio playback**.
- ❑ **Text Extraction:** Tools to process **PDFs, DOCX files, and images** into machine-readable text.
- ❑ **Summarization & Refinement:** AI models that convert **complex legal text into simple summaries**, with refinement for **spoken-friendly outputs**.
- ❑ **Q&A System:** Embedding + generative models for **clause-specific queries** and accurate, contextual answers.
- ❑ **Speech Integration:** Speech-to-text for queries and text-to-speech for **audio-based explanations**.
- ❑ **Privacy & Security:** Data masking to remove sensitive information before processing.
- ❑ **Database (RAG):** Vector storage for efficient search and retrieval in Q&A pipelines.
- ❑ This tech stack ensures a **robust, end-to-end pipeline** — from **document upload to secure, simplified, and voice-enabled understanding**.



Estimated implementation cost

1) **Backend Development (Python + FastAPI/Flask):** Minimal cost if using open-source frameworks; primarily **developer time**. No licensing fees required.

2) **Frontend Development (Streamlit):** Open-source and free; costs mainly include **developer hours** for UI layout, document upload interface, Q&A panel, multilingual support, and audio playback.

3) **AI Models:**
- **flan-t5-large, legal-summarizer-bart, FLAN-T5-small, all-MiniLM-L6-v2 embeddings** – open-source, can run locally on free-tier GPUs or CPUs. Minimal cloud costs if deployed on local machines.
- **Gemini LLM (Google Cloud)** – used for **text refinement and STT**, costs are based on **API usage**. For a prototype with limited documents, cost can remain **very low**, e.g., a few dollars per hundred requests.

4) **Text Extraction Tools (pdfplumber, docx2txt, Tesseract OCR):** Free and open-source.

5) **Text-to-Speech (Coqui TTS):** Free for basic usage; local deployment avoids additional cloud fees.

6) **Database / RAG (FAISS / ChromaDB):** Open-source; no additional cost if hosted locally.

7) **Overall Estimate:** For a hackathon prototype, using mostly **local resources + minimal Gemini usage**, the total implementation cost can be kept **under $50–100**, excluding developer time.

Add as per the requirements for the hackathon:

- ❑ **Prototype Readiness:** The solution is fully functional as a **working prototype**, allowing users to interact with the AI assistant by **uploading documents, viewing summaries, asking clause-specific questions, and listening to audio explanations**.
- ❑ **Privacy Compliance:** Includes **data masking** using **Presidio**, ensuring personal and sensitive information is never exposed, meeting **hackathon expectations for security and ethics**.
- ❑ **End-to-End AI Pipeline:** Demonstrates a complete flow from **document upload → text extraction → summarization → refinement → Q&A → voice output**, clearly showing technical feasibility.
- ❑ **Multilingual Support:** Supports multiple languages for input and output, addressing **inclusivity and accessibility**, which are often evaluated in hackathons.
- ❑ **User Experience Focus:** Provides a **simple, intuitive interface** (via Streamlit) with clear options for document upload, multilingual selection, and audio playback, making it **judge-friendly and easy to demo**.
- ❑ **Implementation Feasibility:** Uses **open-source models and tools** for most processes, with **minimal cloud dependency (Gemini LLM for refinement/STT only)**, showing **practicality and low-cost deployment**.
- ❑ **Visual Presentation:** Includes **flow diagrams, architecture sketches, and UI mockups** to make the solution understandable and visually appealing during the pitch.
- ❑ **Scalability Potential:** users can see that the solution can be extended with **more languages, larger document support, and cloud deployment**, highlighting long-term applicability.

Google Cloud H2S

Gen AI Exchange Hackathon

Thank you