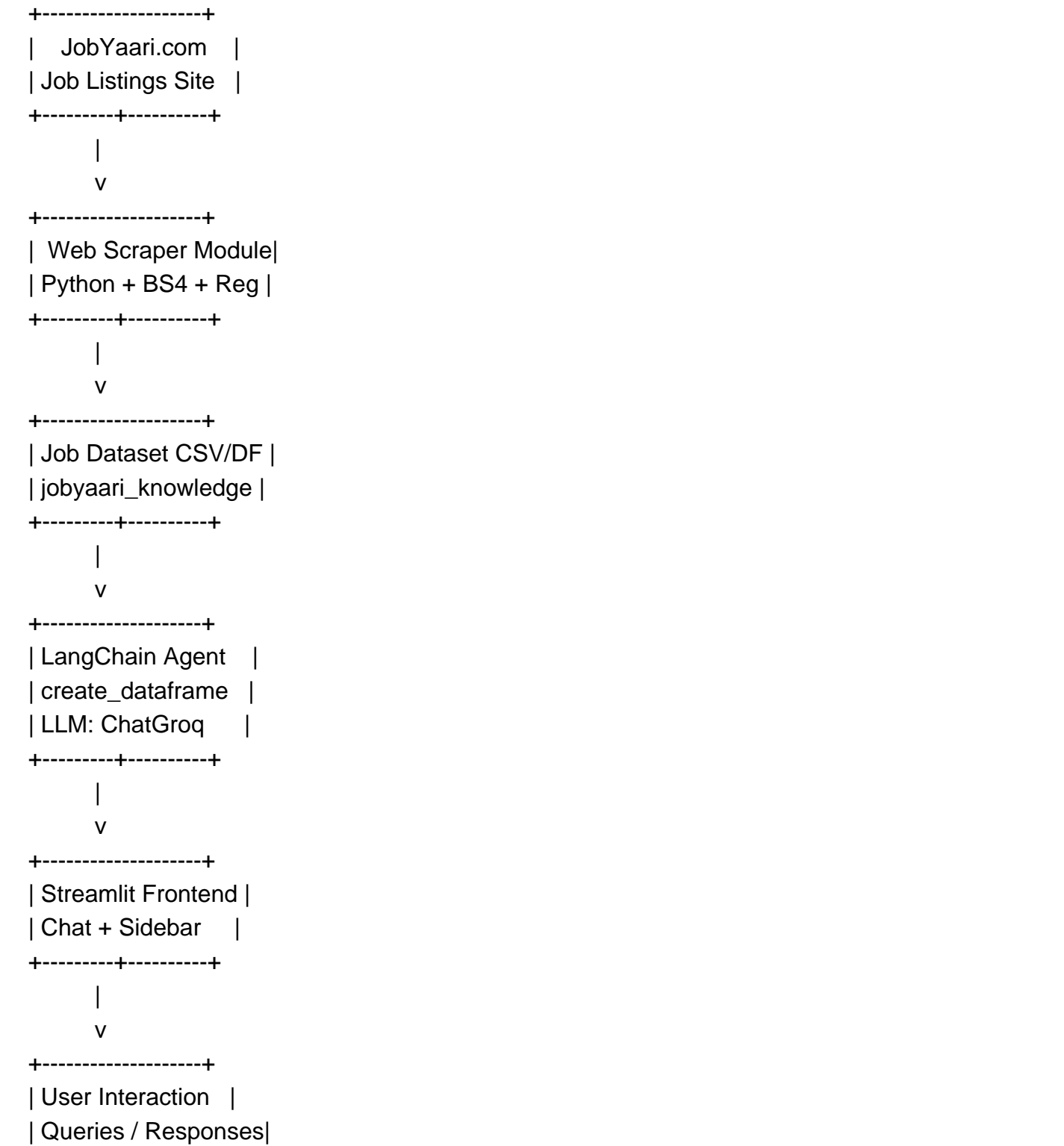# JobYaari AI Assistant - Architecture Blueprint

1. Overview
The JobYaari AI Assistant is a Streamlit-based chatbot that helps users query jobs scraped from JobYaari.com. It leverages web scraping, a Pandas DataFrame Agent, Groq LLM, and a Streamlit frontend.

2. System Architecture Diagram (ASCII version)

```
+-------------------+
|   JobYaari.com    |
| Job Listings Site |
+---------+---------+
          |
          v
+-------------------+
|  Web Scraper Module|
| Python + BS4 + Reg |
+---------+---------+
          |
          v
+-------------------+
| Job Dataset CSV/DF |
| jobyaari_knowledge |
+---------+---------+
          |
          v
+-------------------+
| LangChain Agent   |
| create_dataframe  |
| LLM: ChatGroq     |
+---------+---------+
          |
          v
+-------------------+
| Streamlit Frontend |
| Chat + Sidebar    |
+---------+---------+
          |
          v
+-------------------+
| User Interaction  |
| Queries / Responses|
```

```
    +-------------------+
```

3. Workflow Steps
- Step 1: Data Scraping: Categories -> Fetch job cards -> Extract fields -> Parse URLs -> Fetch details -> Output CSV
- Step 2: Data Storage: CSV & Pandas DataFrame
- Step 3: LLM Integration: LangChain Agent + ChatGroq
- Step 4: Streamlit Frontend: Sidebar refresh, chat interface
- Step 5: User Interaction: Queries processed by LLM, displayed

4. Technology Stack
- Web Scraping: Python, Requests, BeautifulSoup4
- Data Storage: Pandas, CSV
- AI / NLP: LangChain, ChatGroq (Groq API)
- Frontend: Streamlit
- Deployment: Streamlit Cloud / Heroku / Local

5. Integrations
- Groq API, Pandas Agent, Streamlit Session State, Web Scraper

6. Data Flow
1. User opens chatbot -> Streamlit loads dataset
2. User submits query -> LLM queries DataFrame
3. Result -> Displayed in chat interface
4. Optional: Refresh dataset -> Scraper updates CSV

7. Deliverables
- Live Chatbot Link
- Recorded Video Walkthrough
- Blueprint / Architecture Diagram (This PDF)
- Training Data / Knowledge Base CSV
- Google Drive Folder with above items