



BHARATIYA ANTARIKSH HACKATHON

2025

Powered by **H2S**



Team Name : **GravitasOps**

Team Leader Name : Dinesh Yadav

Problem Statement : AI-based Help Bot for Information Retrieval from a Knowledge Graph
Based on Static/Dynamic Web Portal Content

Team Members

Team Leader:

Name: Dinesh Yadav
College: Cluster Innovation Centre,
University of Delhi

Team Member-2:

Name: Anushka Jain
College: Maharaja Agrasen Institute of
Technology, Delhi

Team Member-1:

Name: Sachin Yadav
College: Cluster Innovation Centre,
University of Delhi

Team Member-3:

Name: Aditya Raj
College: Cluster Innovation Centre,
University of Delhi

Brief about the Idea:

An AI-powered conversational assistant that transforms the MOSDAC portal into an intuitive, graph-driven interface. It builds a unified knowledge graph from structured (satellite imagery metadata, geophysical parameters) and unstructured (research papers, technical docs) data, enabling natural-language queries and real-time updates.

Opportunity & Differentiation:

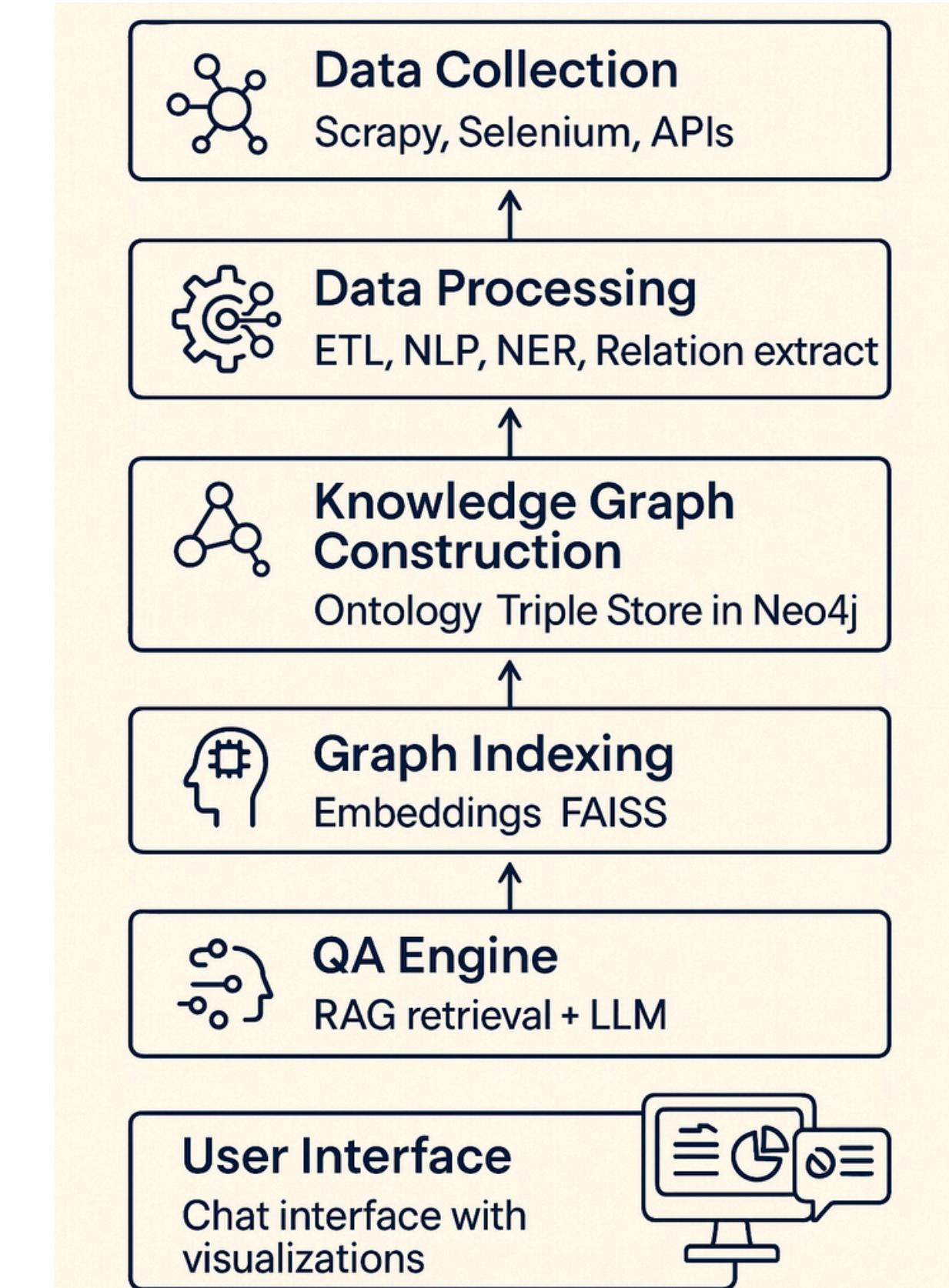
- **Key Differentiator:** Leverages a knowledge graph and conversational AI instead of traditional portals or keyword-based search, enabling natural-language access to MOSDAC's satellite data.
- **Problem Solving Approach:**
 - Unifies real-time and archived data into a semantic knowledge graph
 - Allows plain-language queries with context-aware, explainable answers
 - Bridges the gap between complex data formats and non-technical user access
- **Unique Selling Proposition (USP):**
 - First conversational AI for Indian satellite data
 - Supports real-time updates, multilingual interaction, and interactive visualizations
 - Offers traceable reasoning paths for transparency and trust
 - Empowers a wide range of users—researchers, policymakers, and students

Key Features – MOSDAC AI Knowledge Navigator

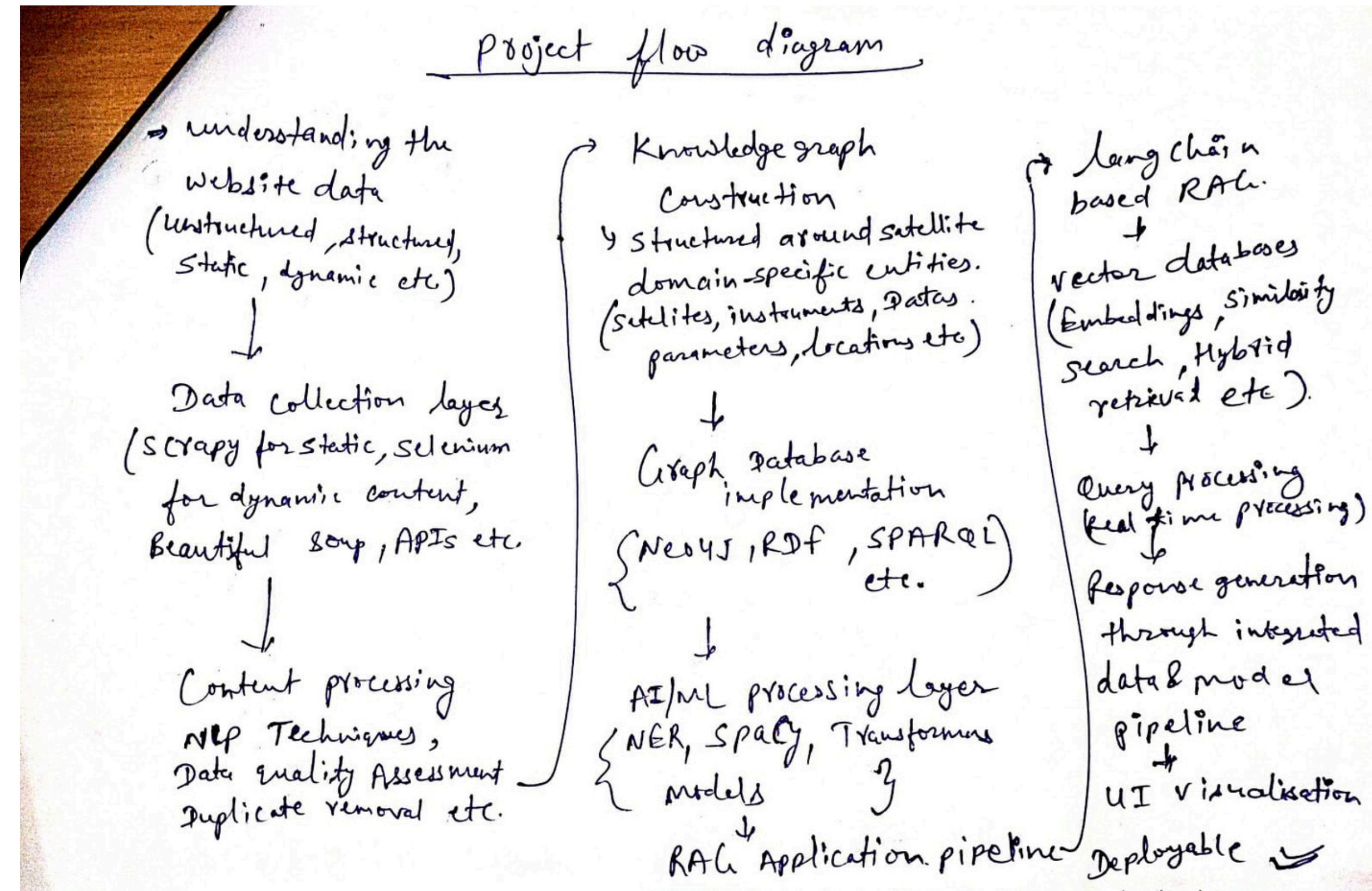
- Natural-language interface for intuitive, plain-language queries
- Knowledge graph backbone integrating structured and unstructured satellite data
- Retrieval-Augmented Generation (RAG) engine for accurate, explainable responses
- Real-time data integration from live MOSDAC feeds
- Interactive visualizations including maps, charts, and graph views
- Explainable reasoning paths showing how each answer is derived
- Contextual follow-up handling for multi-step queries and comparisons
- Multilingual support for both queries and responses
- Multi-modal content delivery with links to imagery, data files, and documents
- Custom alerts and notifications based on user-defined conditions
- Role-based access control for secure and personalized data access
- Extensible plugin framework for adding new data sources and tools

Process flow diagram

- Ingest: Auto-fetch MOSDAC content via APIs and scrapers
- Preprocess: Clean and normalize structured and unstructured data
- Extract: Use NLP to generate triples and build a knowledge graph
- Embed: Vectorize entities for semantic search
- Respond: RAG engine retrieves context; LLM generates answer
- Display: UI shows responses with charts and graphs
- Update: Live data continuously refreshes the graph

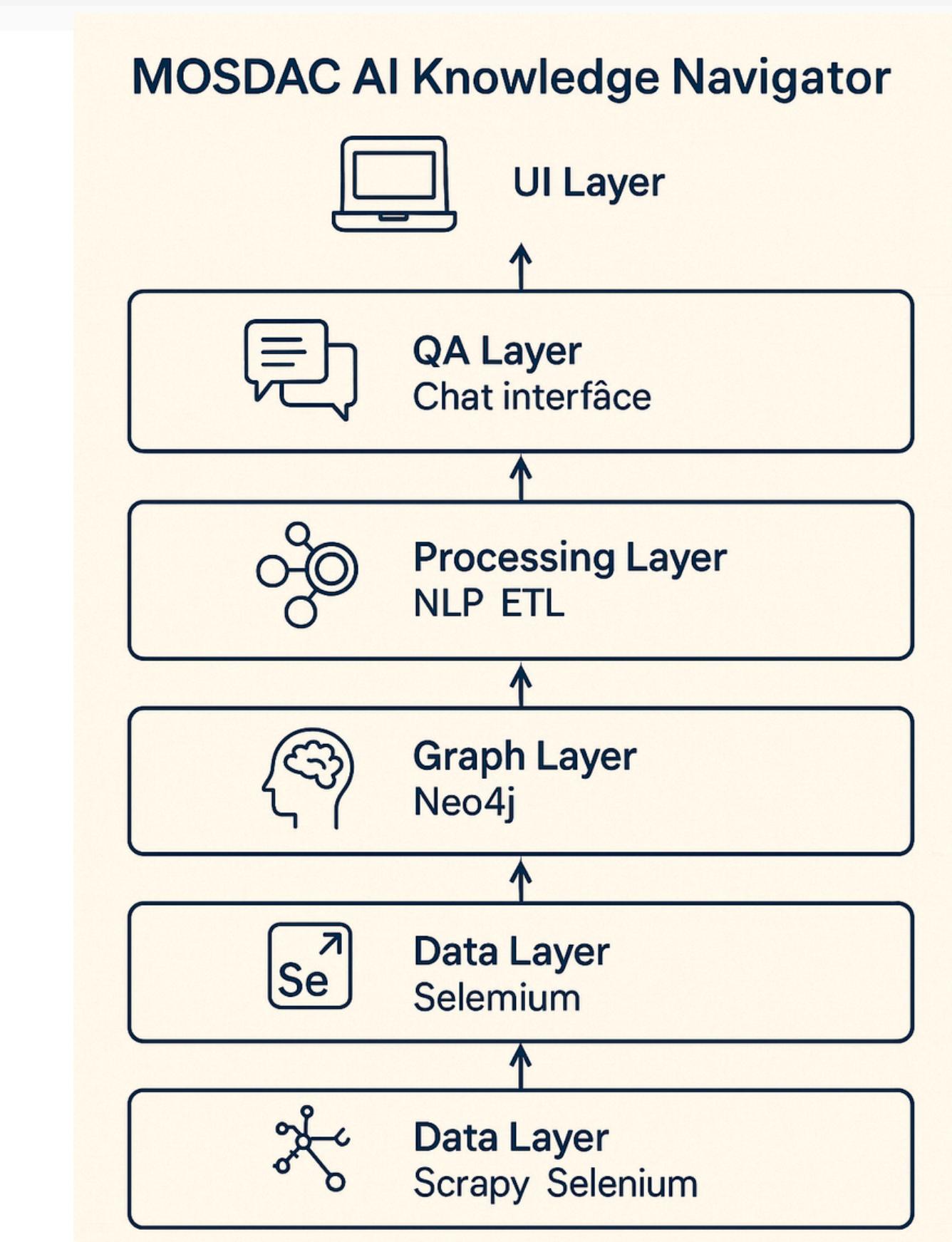
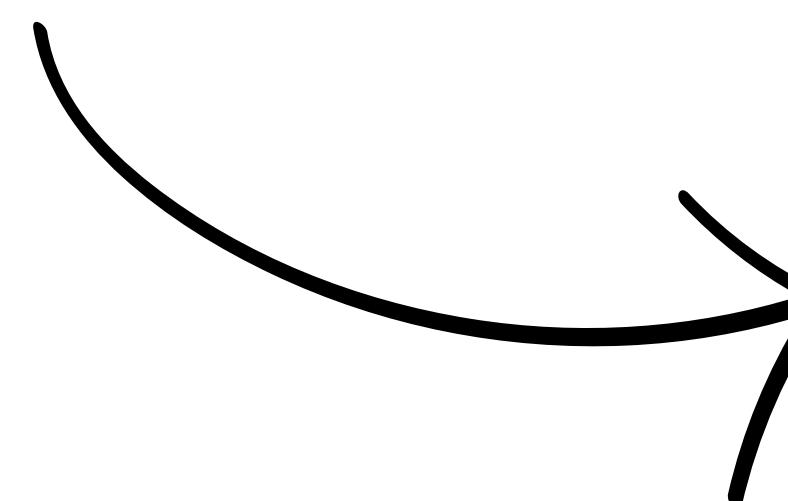


Wireframes/Mock diagrams of the proposed solution:



Architecture diagram of the proposed solution :

Shows how MOSDAC data is ingested, processed into a knowledge graph, and served via QA API



Technologies to be used in the solution:

- AI/ML: GPT-4 (LLM) {or any other LLM which suits us cost and usability-wise}, SciBERT, spaCy NER, LangChain RAG
- Graph DB: Neo4j + RDFLib + SPARQL endpoint
- Vector Store: FAISS/Pinecone for embeddings
- Data Ingestion: Scrapy, Selenium, Airflow ETL
- Backend: FastAPI, PostgreSQL, Redis, RabbitMQ
- Frontend: React+TypeScript, D3.js, WebSocket
- Infra: Docker, Kubernetes on AWS/GCP, Prometheus/Grafana

Expected Outcomes & Impact:

- 80% reduction in time spent on data discovery
- Broadens MOSDAC usage to educators, policymakers, students
- Accelerates research by providing contextual, explainable answers
- Supports ISRO's mission to democratize satellite data

Cost Estimate

- **Primary cost drivers:**
 - Usage of Large Language Models (LLMs) based on token consumption
 - API access for external services (e.g., translation, visualization, alerts)
 -
- **Estimated cost for this project:** ₹3,000 – ₹8,000
 - Optimized through selective querying and efficient model usage
 - Scalable based on actual demand and user volume



Powered by **HACK2SKILL**

BH RATIYA NTARIKSH HACKATHON

2025

THANK YOU

