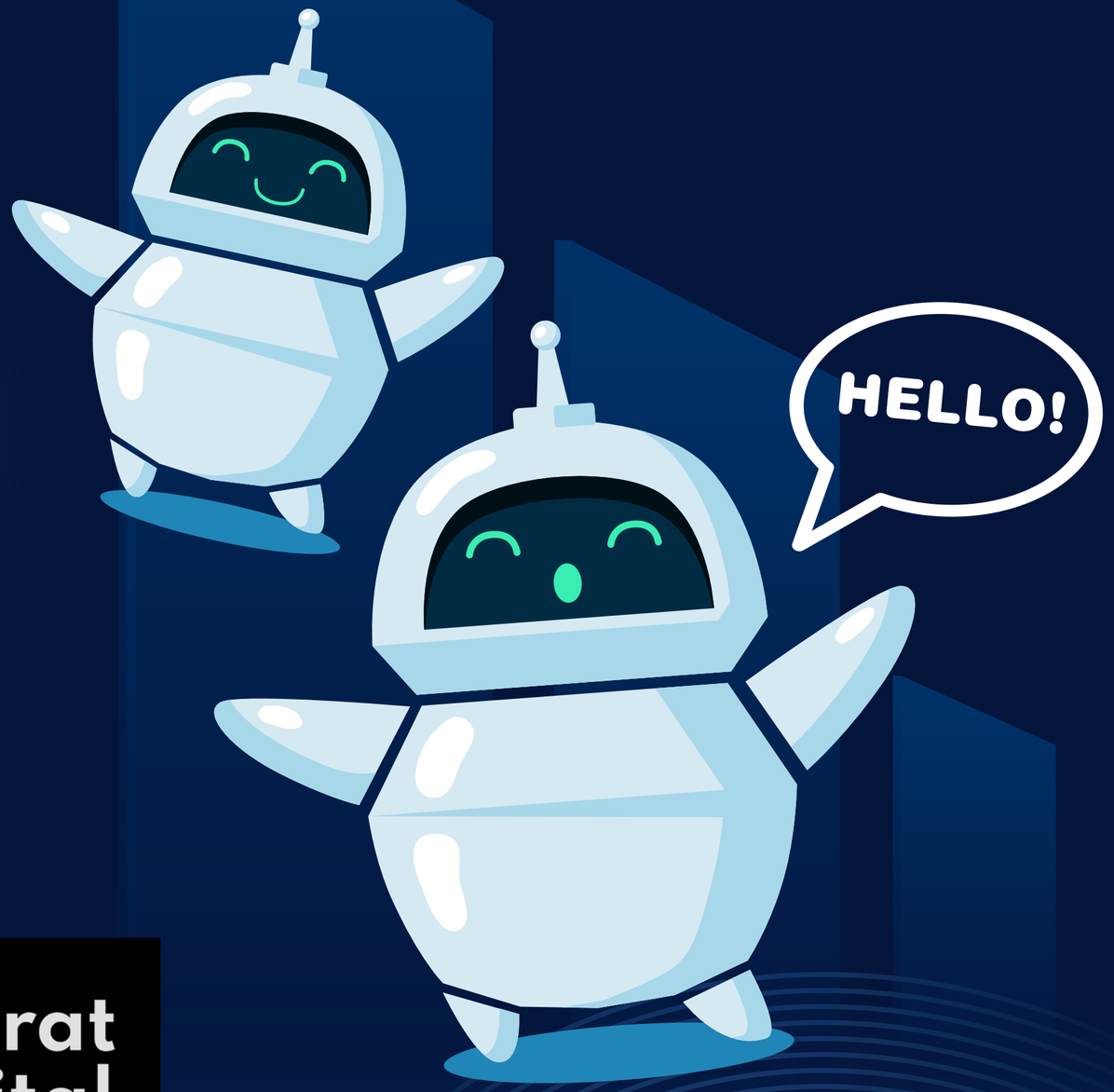
Project Samarth

Intelligent Q&A System for Agriculture and Climate Data



Bharat Digital

Data.gov.in

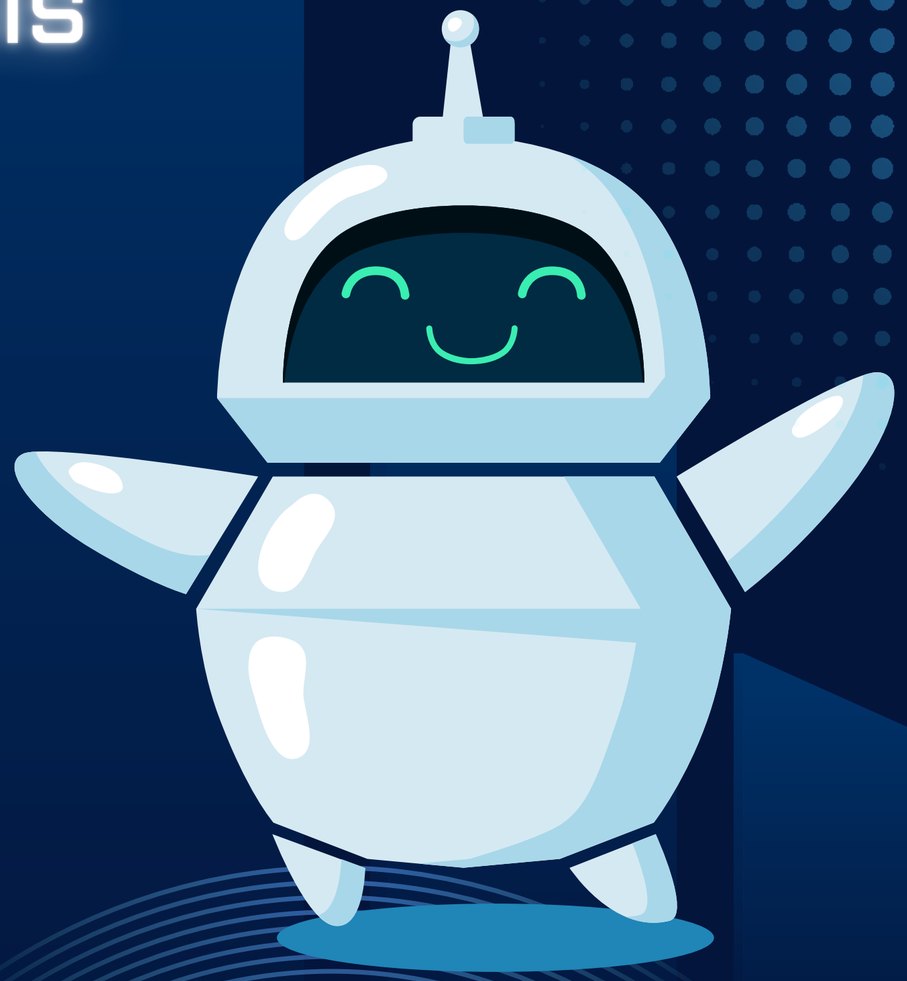
Digital Bharart

“Hello, I’m Ashik Khan, and this is my project Samarth.

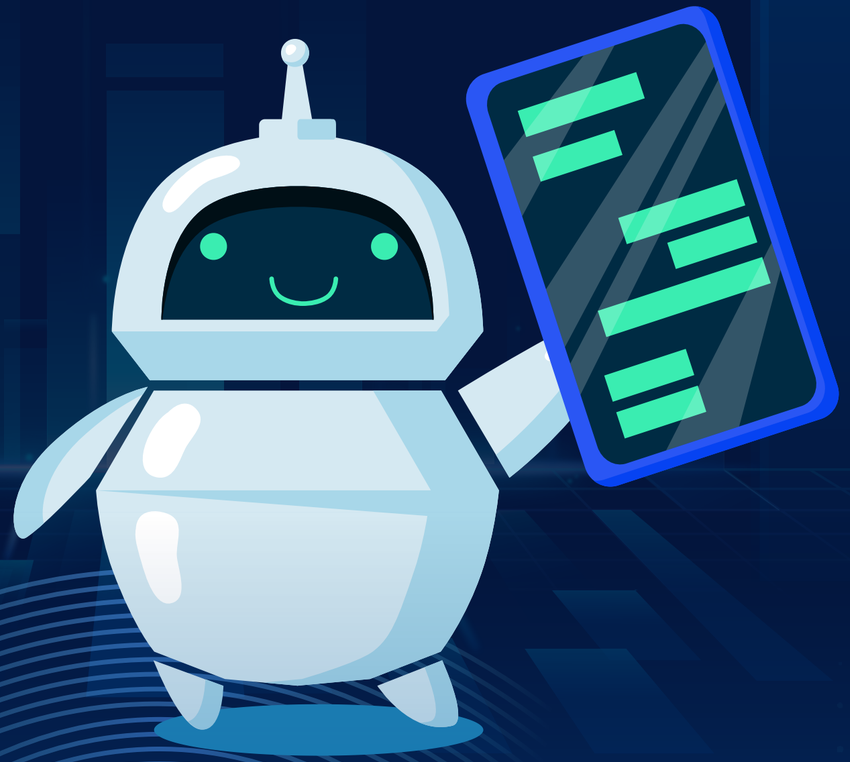
The goal of this project is to make government open data from data.gov.in easier to use.

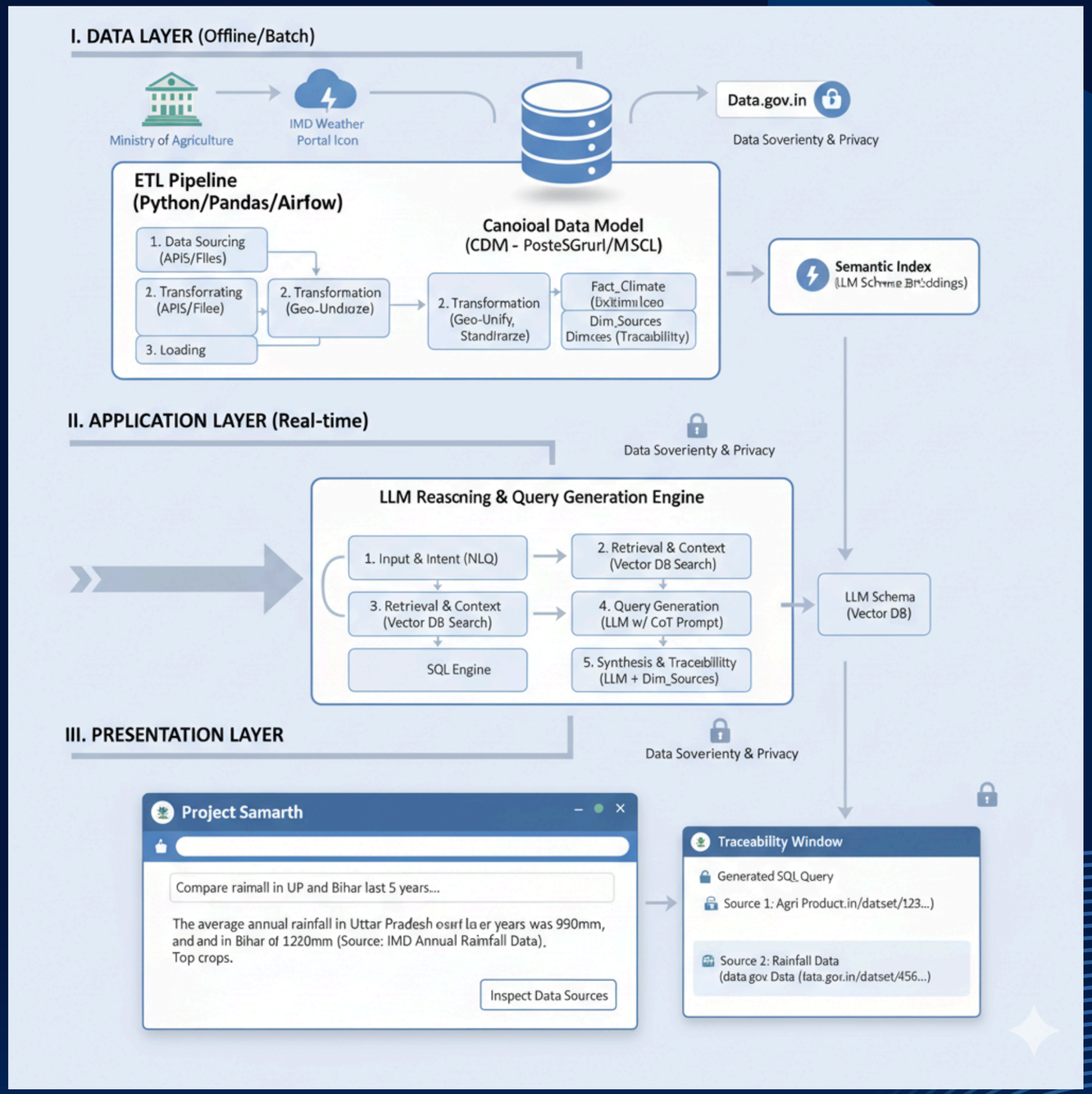
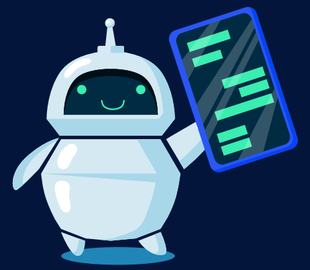
Right now, data from ministries like Agriculture and IMD exist in different formats, which makes analysis hard.

Samarth solves this by building an intelligent Q&A chatbot that connects directly to these datasets and answers complex questions about agriculture and climate.”



System Design of The Project Samarth




I. DATA LAYER (Offline/Batch)
Ministry of Agriculture
IMD Weather Portal Icon
Data.gov.in
Data Soverienty & Privacy
ETL Pipeline (Python/Pandas/Airfow)
1. Data Sourcing (APIS/FIles)
2. Transforrating (APIS/Filee)
3. Loading
2. Transformation (Geo-Undiaze)
Canoioal Data Model (CDM - PosteSGrurl/MSCL)
2. Transformation (Geo-Unify, Standlrarze)
Fact_Climate
Dim_Sources
Dimcees (Tracaibilliity)
Semantic Index
II. APPLICATION LAYER (Real-time)
Data Soverienty & Privacy
LLM Reascning & Query Generation Engine
1. Input & Intent (NLQ)
2. Retrieval & Context (Vector DB Search)
3. Retrieval & Context (Vector DB Search)
4. Query Generation (LLM w/ CoT Prompt)
SQL Engine
5. Synthesis & Traceibillity (LLM + Dim_Sources)
LLM Schema (Vector DB)
III. PRESENTATION LAYER
Data Soverienty & Privacy
Project Samarth
Compare raimall in UP and Bihar last 5 years...
The average annual rainfall in Uttar Pradesh osıf la er years was 990mm, and and in Bihar of 1220mm (Source: IMD Annual Raimfall Data).
Top crops.
Inspect Data Sources
Traceability Window
Generated SQL Query
Source 1: Agri Product.in/datset/123...)
Source 2: Rainfall Data (data gov Dsta (fata.gor.in/datset/456...)