# **Challenge: Solar Detective: Mapping India’s Solar Infrastructure Using Agentic AI**

|  |  |
| --- | --- |
| **Team:** Agrobotsz | **Members: Lakshya , Karthik raja , jeeva** |

## **Data Collection**

* **Agents and Tools**: Using Selenium with a Chrome web driver, intelligent agents crawl various government and commercial websites.
* **Extraction Targets**: The agents identify the following raw datasets:
  + *Company Projects Data:* These are projects from major companies (e.g. TATA Power, Azure Power, ReNew). We collect location, commission date, size, etc.
  + *Government Tender Data:* There are available tenders offers from the government (i.e. SECI). We collect location, tender details, URL for tender documents and submission, etc.
  + *National Statistics Data:* This includes information about climate, population & demographics, and economic development for the different states in India.

## **Data Processing**

* **Parsing and Cleaning:** BeautifulSoup is used for HTML parsing; extracted data is cleaned and structured into CSVs.
* **Geolocation:** GeoPy, with OpenStreetMap backend, converts textual location data into geographical coordinates.
* **Transformation:** Python scripts standardize formats and remove duplicates, while tools like Table Studio and Excel assist in manual verification and annotation when needed.

## **Data Aggregation**

* CSV files are aggregated into structured datasets and stored locally.
* PostgreSQL can serve as a central database for querying and linking spatial data. This to-be-added feature will allow for online updates.

## **Visualization and Deployment**

* **Mapping**: Folium generates interactive maps of solar sites.
* **Interface**: Streamlit powers the web dashboard, allowing users to filter and explore *existing solar installations* across India; as well *available tender offers* and *investment potentials*. This addresses the needs for different stakeholders; *competing energy companies*, *grid operators* and *planners*.