5 ways to use generative AI in geospatial: A practical roadmap

CARTO

Matthew Forrest

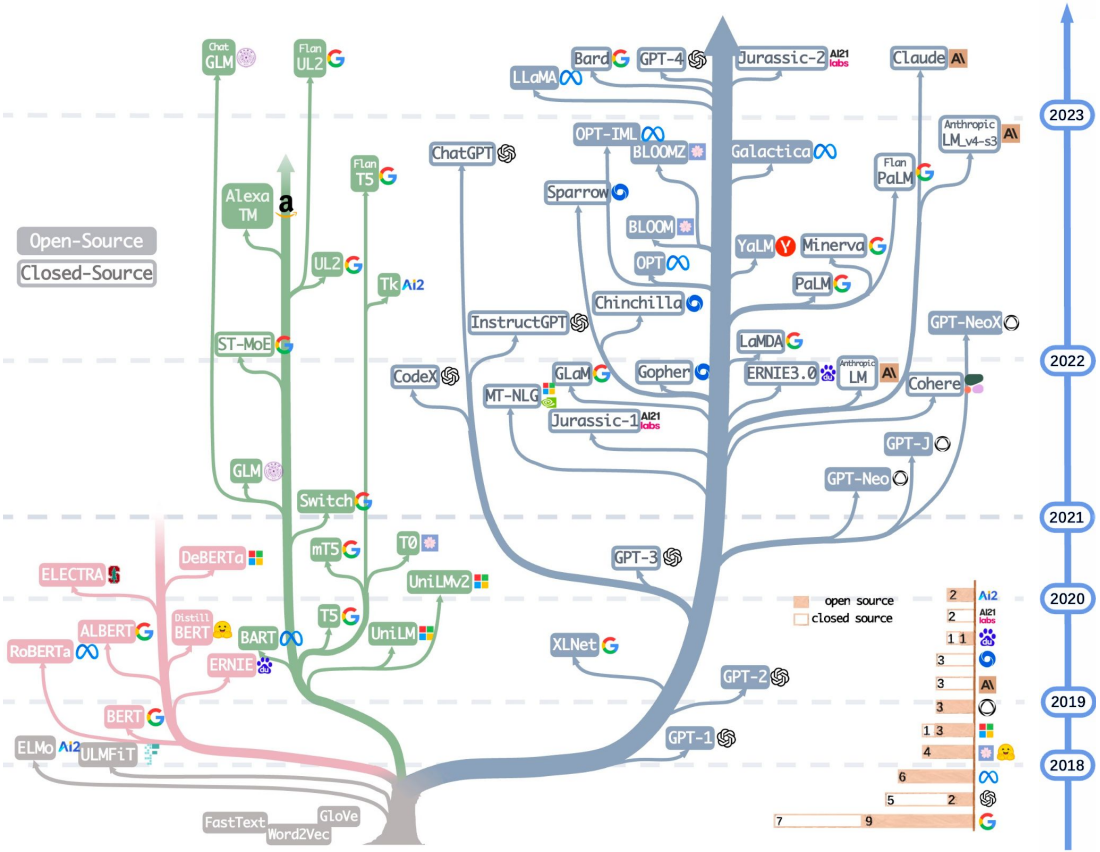
Field CTO

CARTO

The Five Steps to Geospatial generative AI

Code Generation	Writing code to assist with human driven processes
Feature Labeling	Making geospatial features easy to understand
Retrieving Data	Getting data from human based text
Interaction	Use text to manipulate your data or application
Building Custom Models	Building specific models for geospatial

LLMs have come a long way



https://github.com/Mooler0410/LLMsPracticalGuide

Improving our code and writing it faster

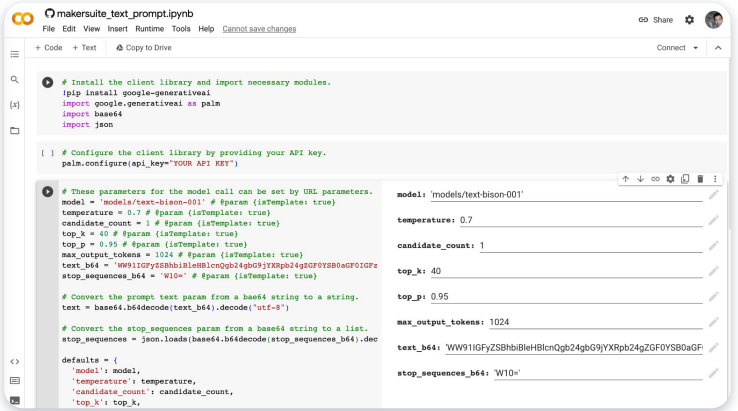
Already available for analytics

In SQL

Query Editor

```sql
SELECT ml_generate_text_result
FROM ML.GENERATE_TEXT(
   MODEL `cartobq.sdsc23_demos.llm_model_sdsc_demo`,
   (SELECT 'Given a set of parameters ... [0.9…]' as prompt))
```

In Python



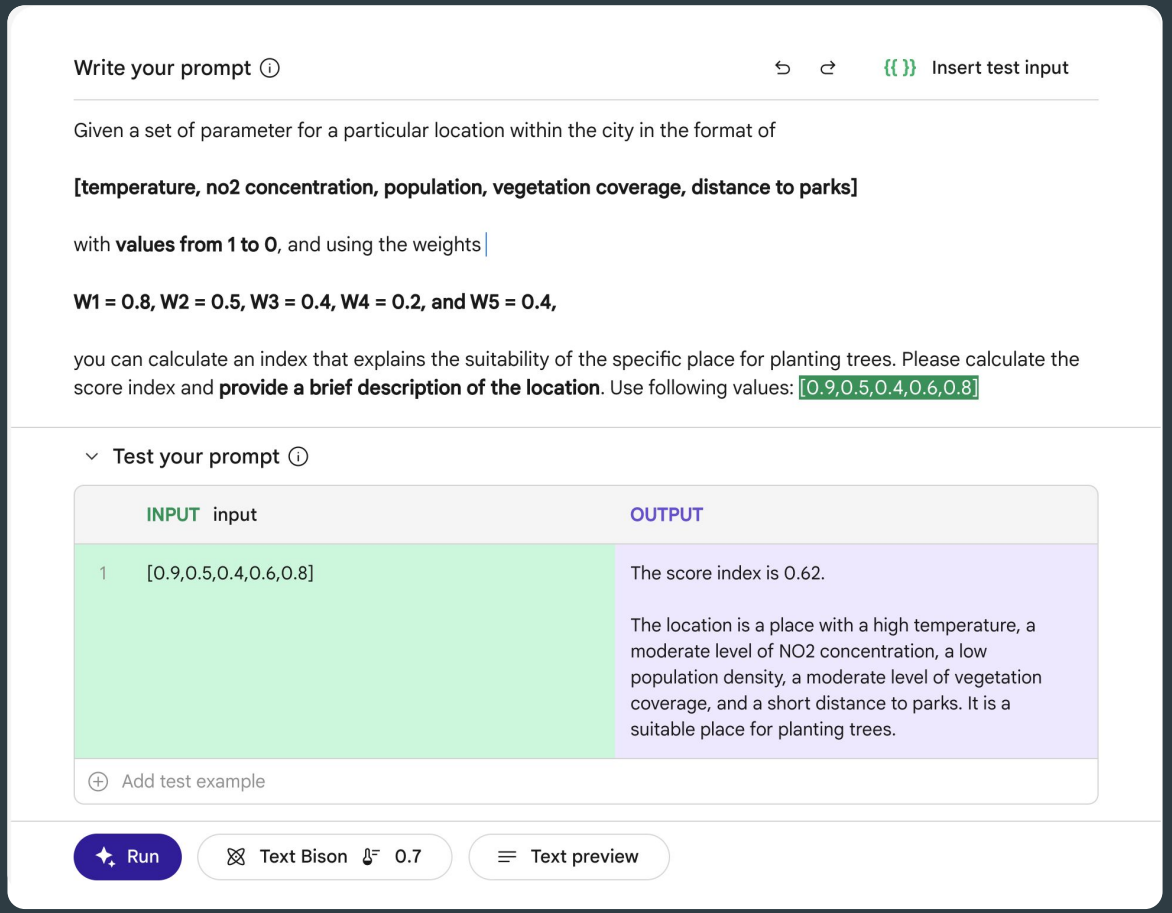
curl:// JS …

Let’s see examples!

Making complex data easier to use

How do you generate 5,000 different labels for each building/score?

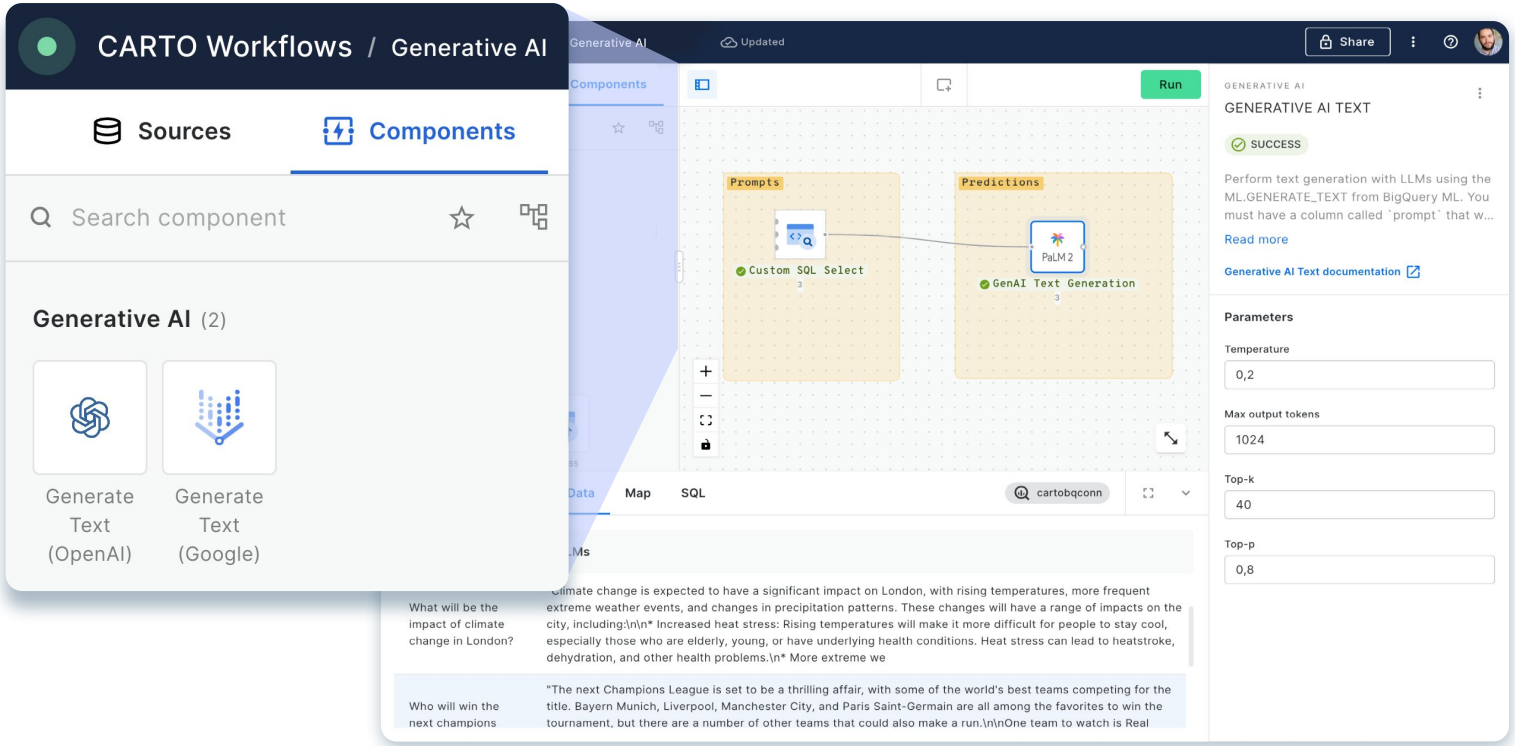
Automatic generation of labels for model results

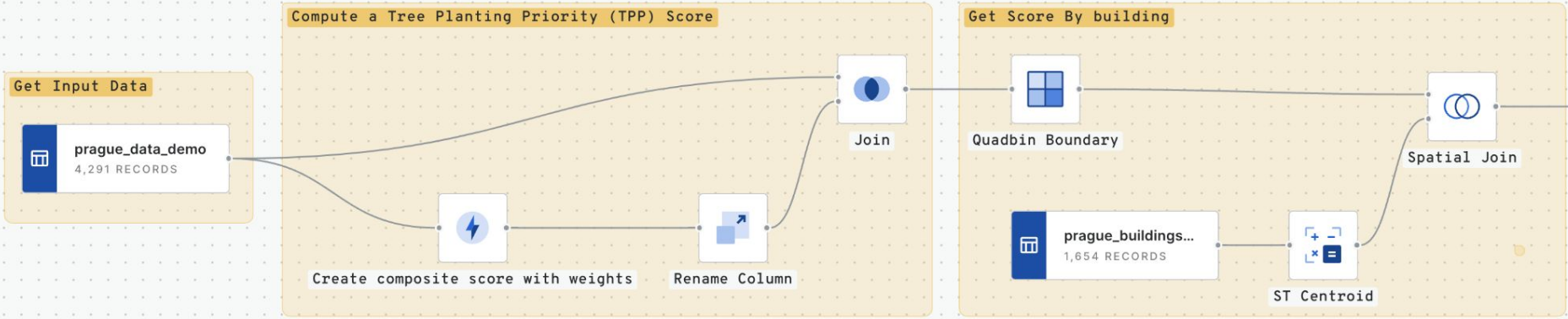


Air Quality, House Price, Crime rate, Amenities and School Quality

This location boasts excellent air quality, high house prices, and a low crime rate, making it suitable for those prioritizing safety and a healthy environment. However, it has limited amenities and poor school quality, which may deter families seeking educational opportunities and a variety of amenities nearby.

In CARTO Workflows!





Create Score Index

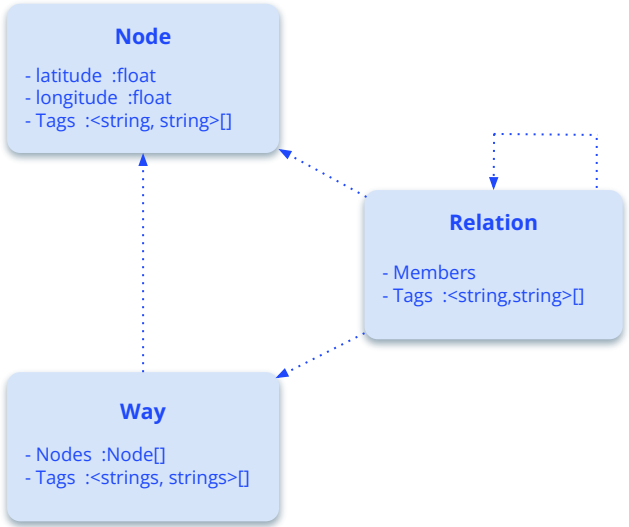
Bring it to buildings


Automatic Segmentation of Trees
Not just for vector data!

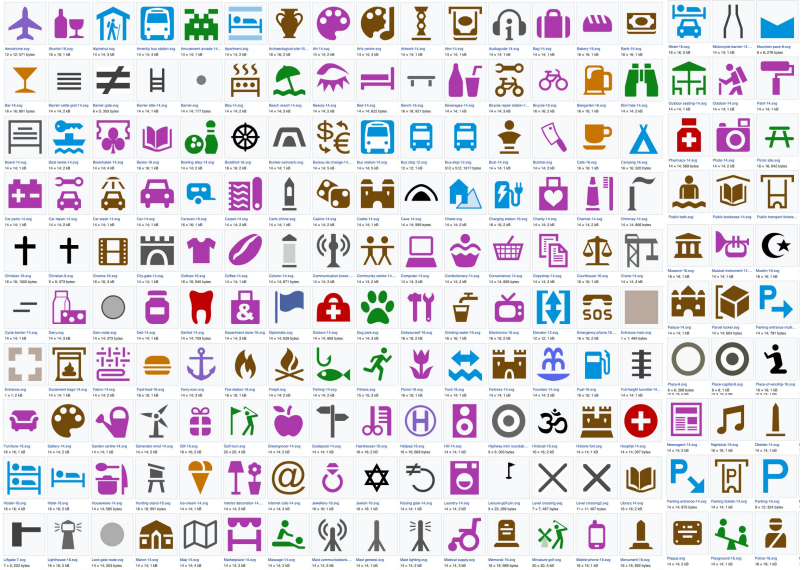
Using text to get our data for us

Writing SQL for OpenStreetMaps can be painful

All data already available in different platforms



Schema-less



k=v 92K tags for 3B objects

Generation of complex Spatial SQL

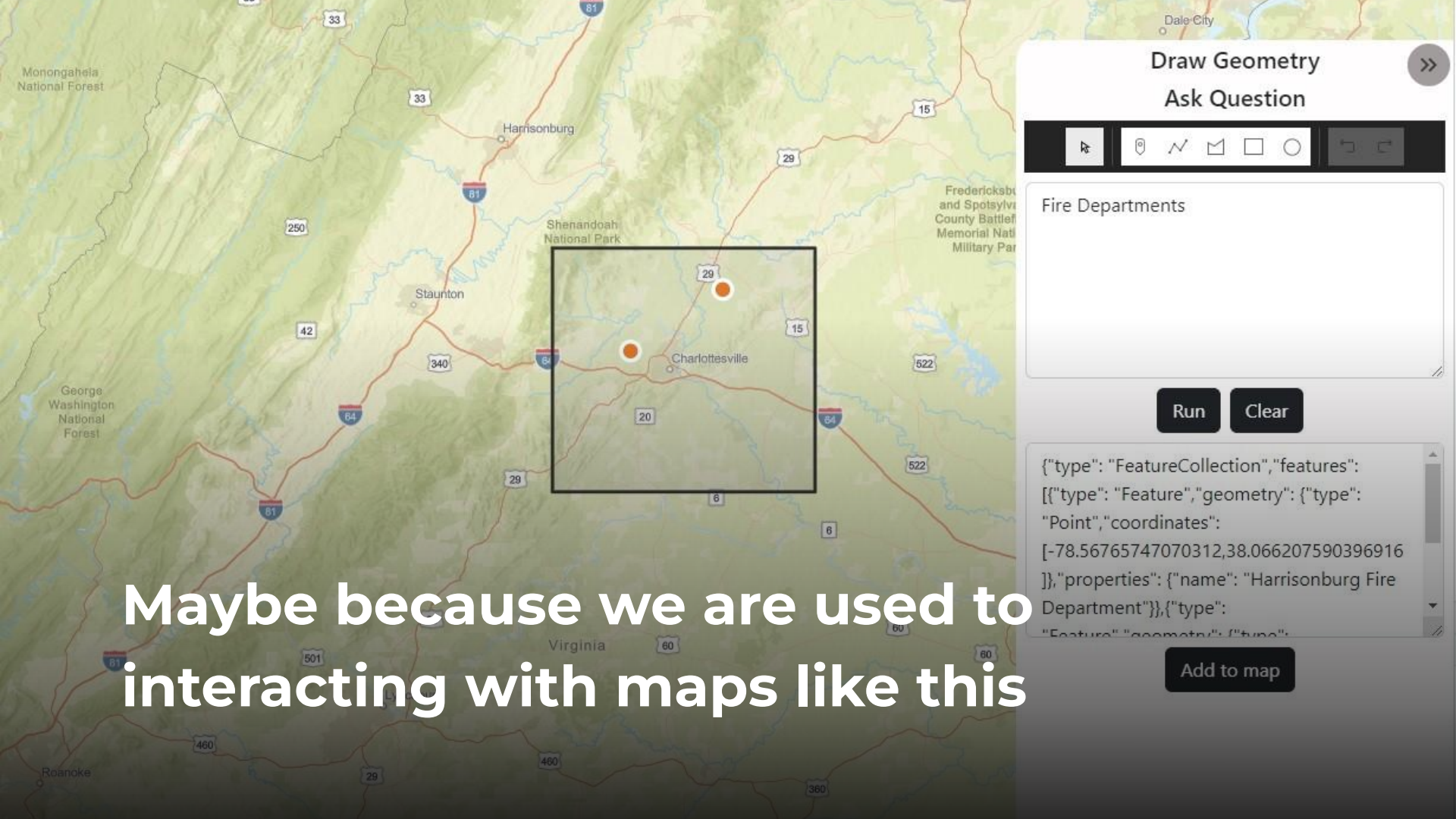
Give me all healthcare facilities in South of Spain

If you want to find all point features with any value for the 'healthcare' tag in the southern part of Spain, here's the SQL query:

sql Copy code

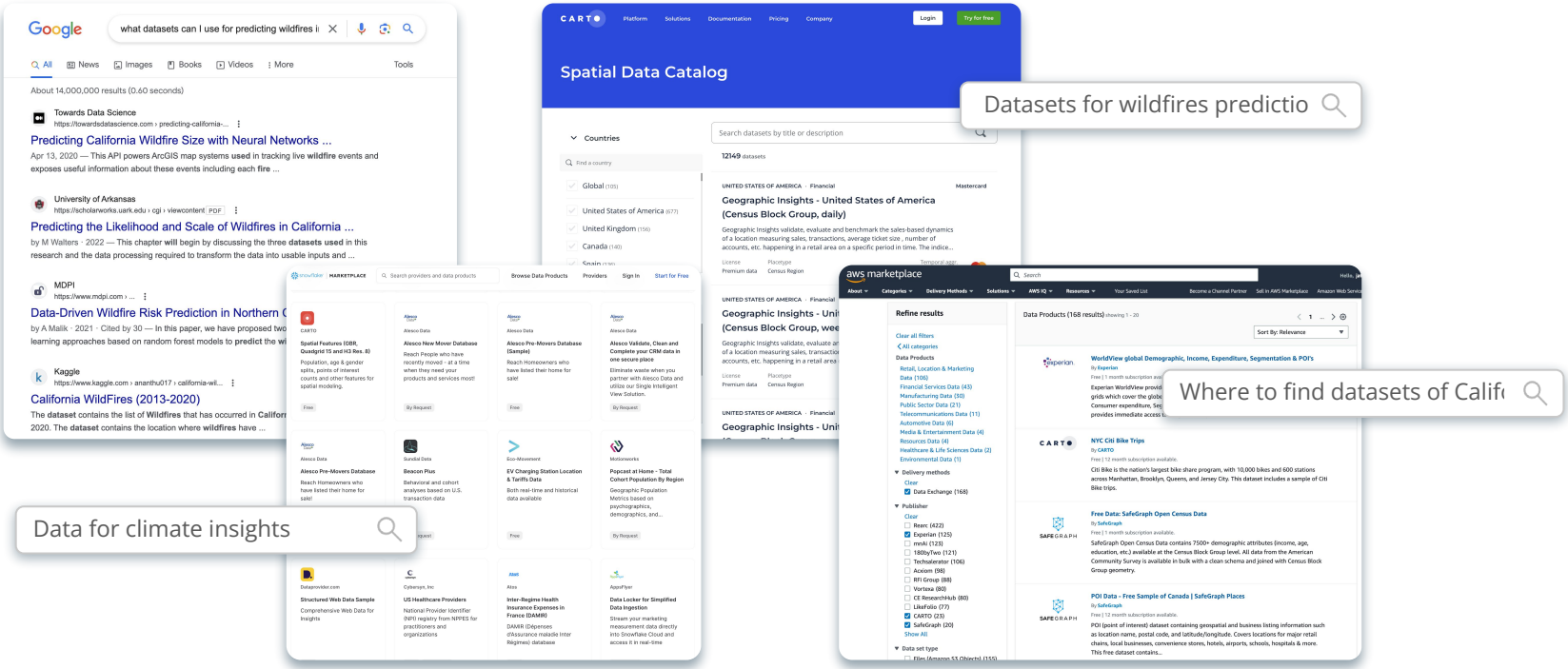
```sql
SELECT
    feature_type,
    osm_id,
    osm_timestamp,
    geometry AS geom,
    (SELECT value FROM UNNEST(all_tags) WHERE key = '
FROM `bigquery-public-data.geo_openstreetmap.planet_f
WHERE feature_type = 'points'
AND EXISTS(SELECT 1 FROM UNNEST(all_tags) WHERE key =
AND ST_INTERSECTS(geometry, ST_GEOGFROMTEXT('POLYGON(
LIMIT 1000;
```


CARTO Builder / Untitled
Updated
Share
Map view
Layers
TILES
healthcare in South Spain
Search location
Córdoba
Alicante
Portimão
Huelva
Sevi
Faro
ANDALUSIA
Granada
Cartagena
Cádiz
Málaga
Almería
GIBRALTAR
© CARTO, © OpenStreetMap contributors
SQL Editor - Edited
carto_dw
1 SELECT carto.us.QUERY_OSM('Give me healtchare facilities in South of Spain')
SOURCES (1)
Query · BigQuery
SQL Query 2
Add source from...
Run
00:05
Create table from query

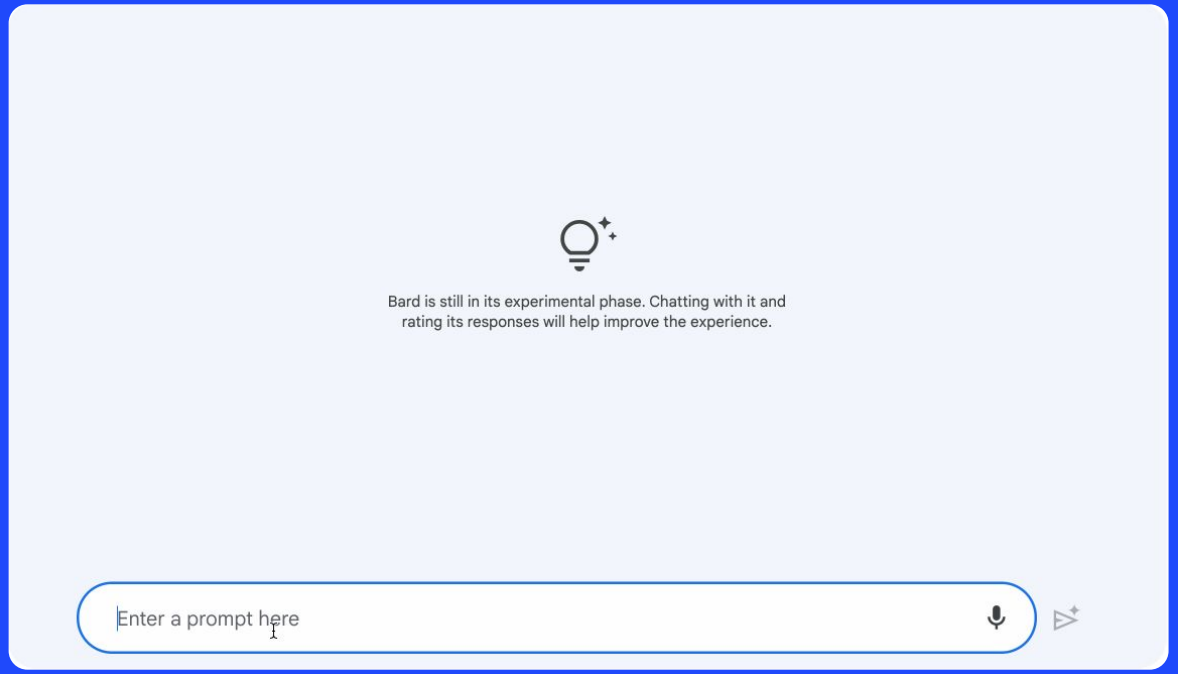

Draw Geometry
Ask Question
Fire Departments
Run
Clear
{"type": "FeatureCollection","features": [{"type": "Feature","geometry": {"type": "Point","coordinates": [-78.56765747070312,38.066207590396916]},"properties": {"name": "Harrisonburg Fire Department"}},{"type":
Add to map
Maybe because we are used to interacting with maps like this
Harrisonburg
Staunton
Charlottesville
Shenandoah National Park
Monongahela National Forest
George Washington National Forest
Virginia
Roanoke
Dale City

Find datasets relevant to your analysis

More than 50% of time is often spent searching data

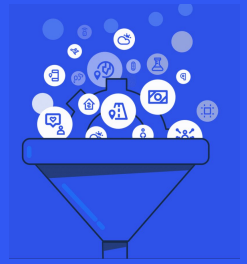


Your personalized expert suggesting datasets



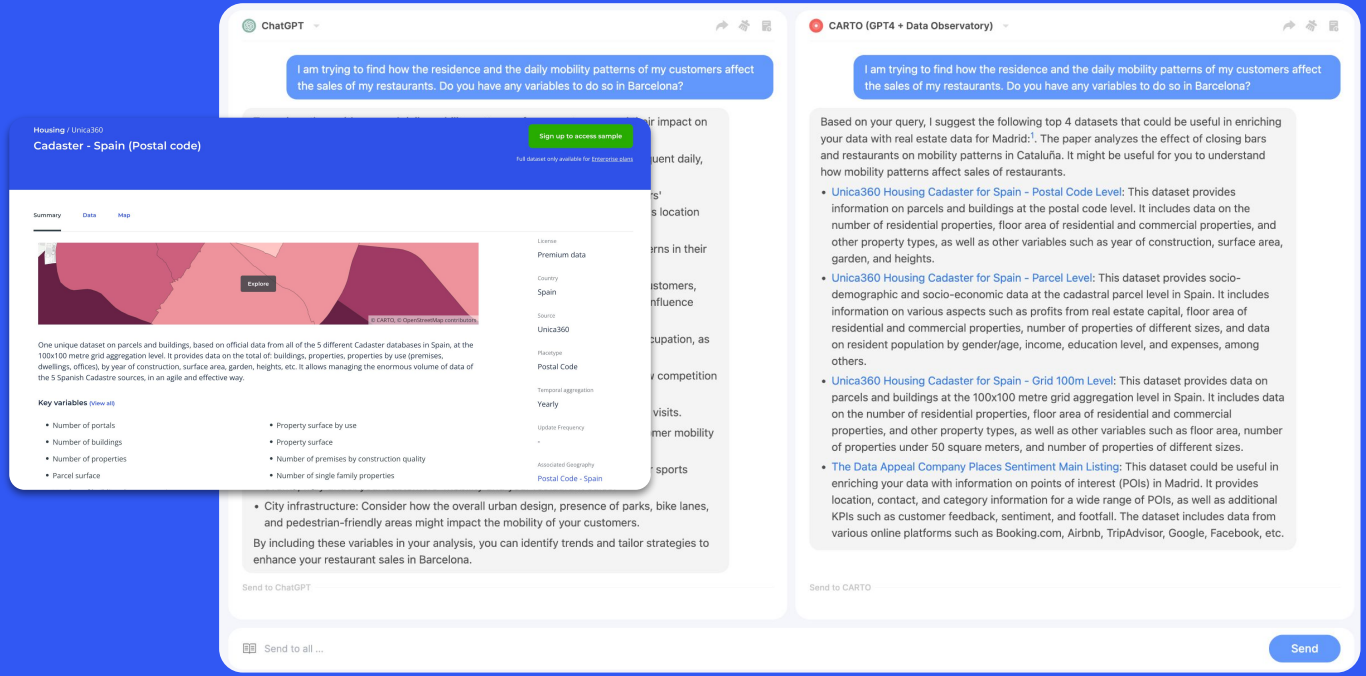
Finding data relevant to your analysis

Fine-tuning the model with CARTO Data Observatory knowledge



12,149 datasets
120 countries
40 providers
8 categories

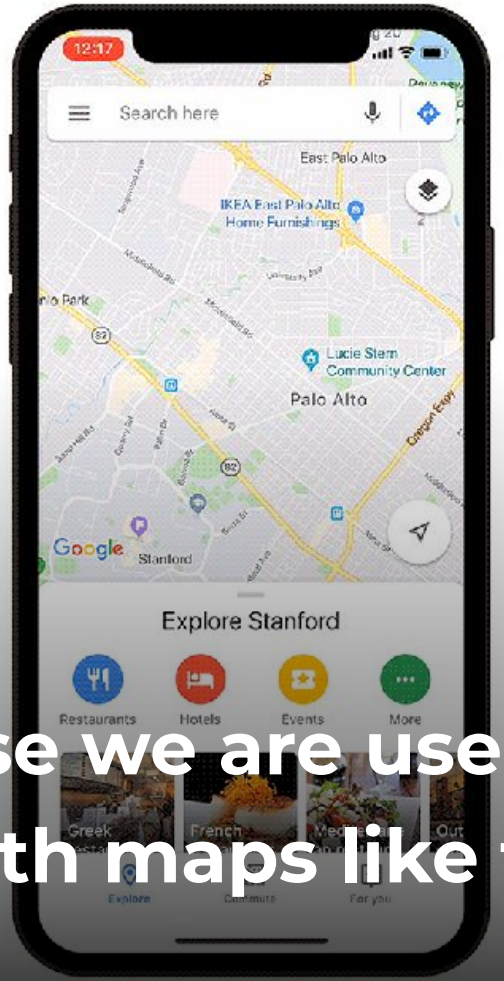
CARTO



Interacting with our maps

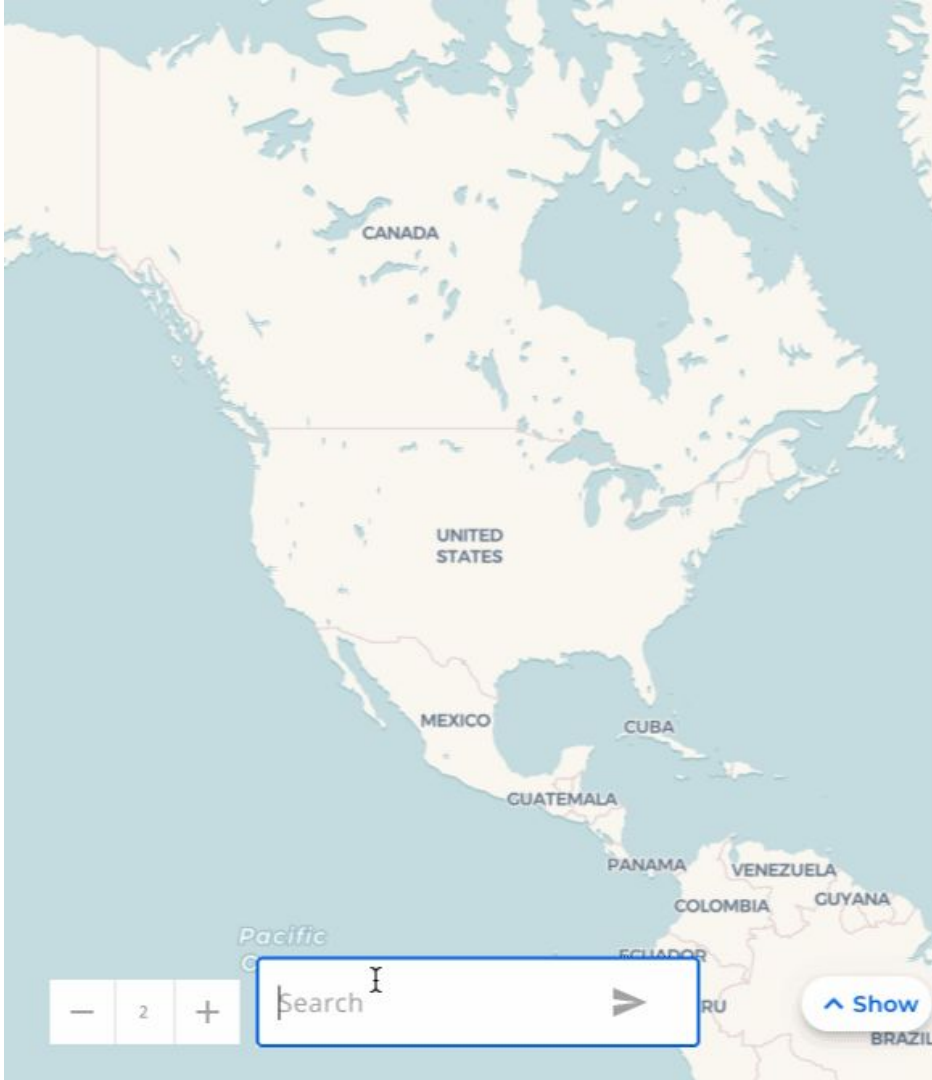
AI is the new UI

And SQL remains the lingua franca for Analytics, now on a new AI world of LLMs



Maybe because we are used to interacting with maps like this

- ✍️ Prompt Engineering
- 📓 Embedding CARTO DO
- ⚙️ Spatial SQL in LLMs
- 🗺️ CARTO Maps API
- 🌐 **MapGPT!**

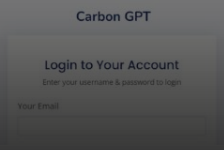


Revolutionising Carbon Management With AI

JOIN WAITLIST LEARN MORE

AI-driven, cloud-based carbon footprint management platform designed for organizations

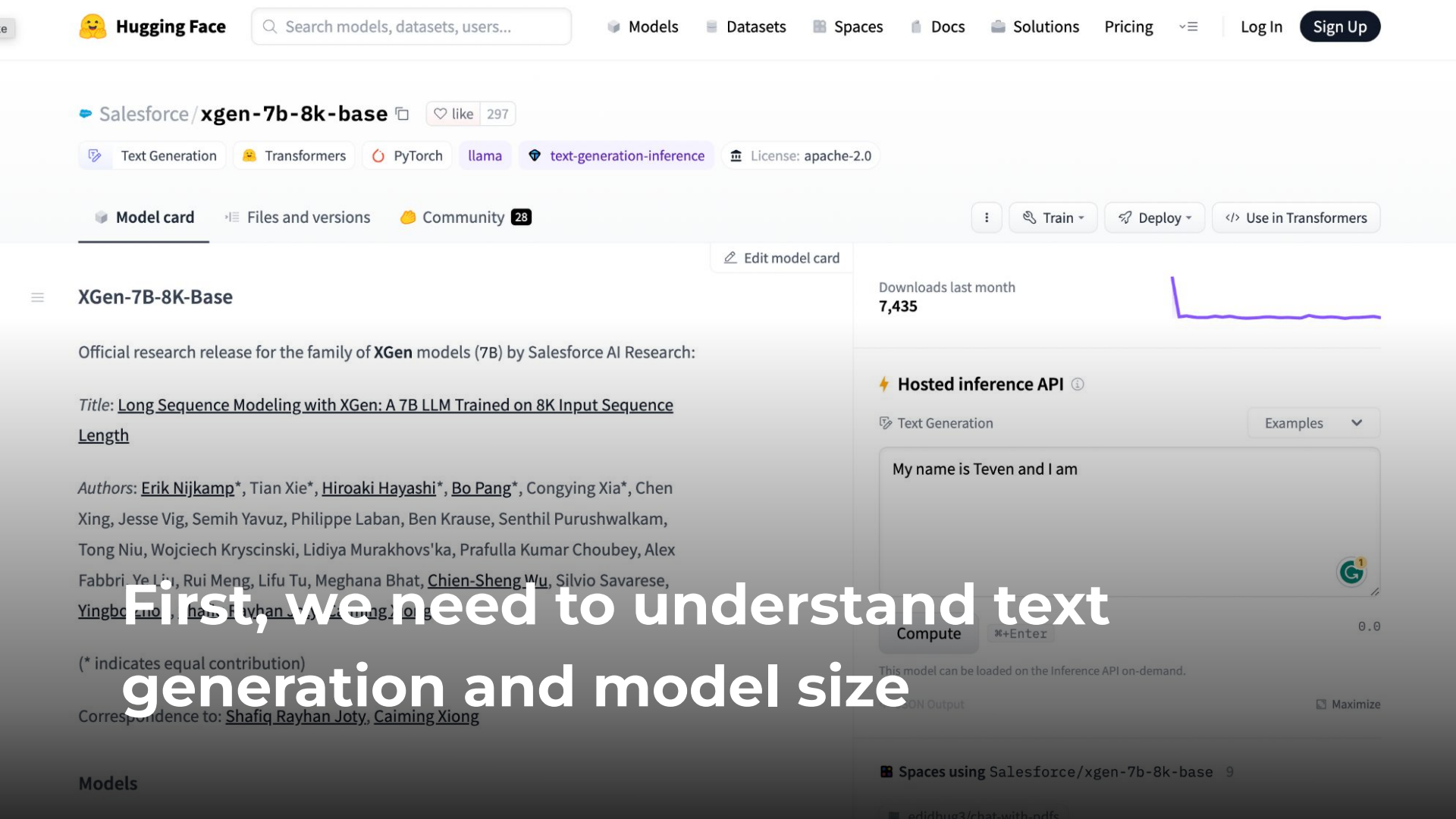
Connecting text based datasets to geospatial data




So where is
this going?

Custom models for geospatial

First, what are these models good for?



First, we need to understand text generation and model size

Improving Machine Learning from Human Feedback

Erin Mikail Staples + Nikolai Lubimov

PyData DE 2023

Hugging Face | Search models, datasets, users... | Models | Datasets | Spaces | Docs | Solutions | Pricing | Log In | Sign Up

IBM NASA Geospatial Community

Research interests

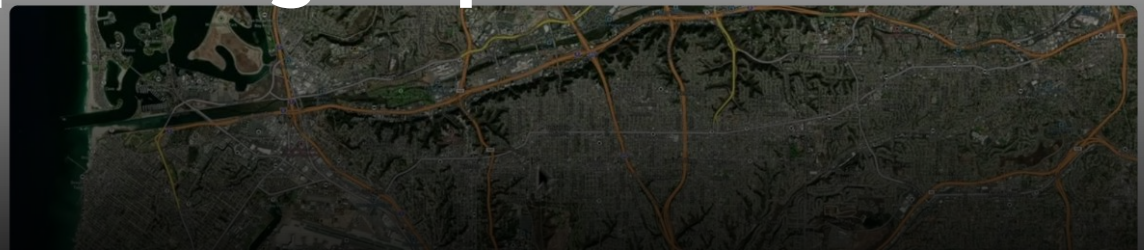
Geospatial foundation models using HLS2 data

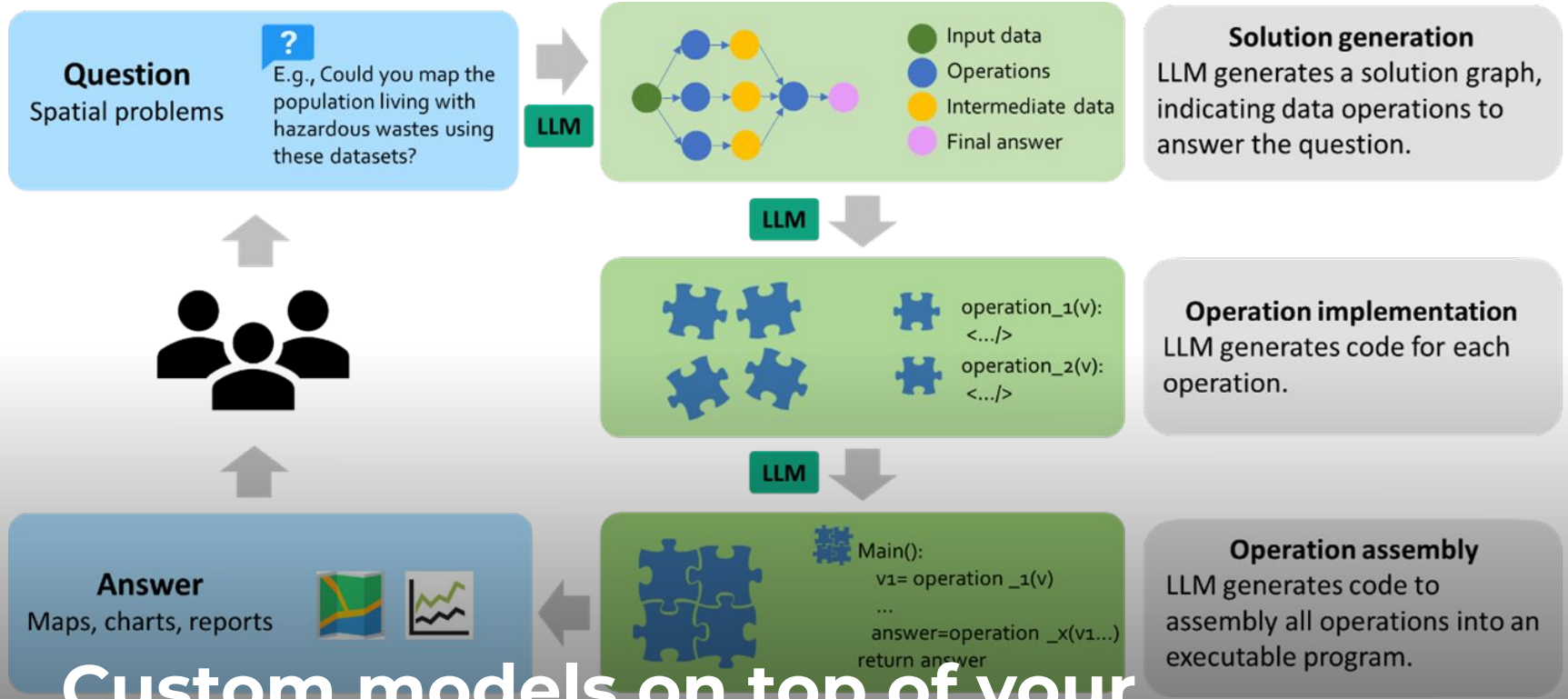
Team members 38

Organization Card About org cards

NASA and IBM have teamed up to create an AI Foundation Model for Earth Observations, using large-scale satellite and remote sensing data, including the Harmonized Landsat and Sentinel-2 (HLS) data. By embracing the principles of open AI and open science, both organizations are actively contributing to the global mission of promoting knowledge sharing and accelerating innovations in addressing critical environmental challenges. With Hugging Face's platform, they simplify geospatial model training and deployment, making it accessible for open science users, startups, and enterprises on multi-cloud AI platforms like watsonx. Additionally, Hugging Face enables easy sharing of the pipelines of the model family, which our team calls Prithvi, within the community, fostering global collaboration and engagement.

Models for specific geospatial use cases





Custom models on top of your data