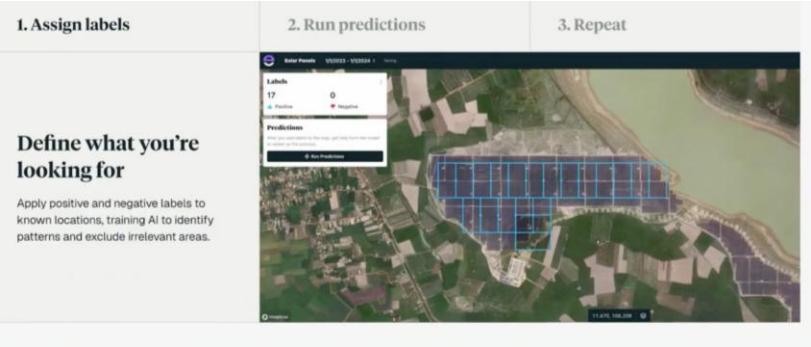
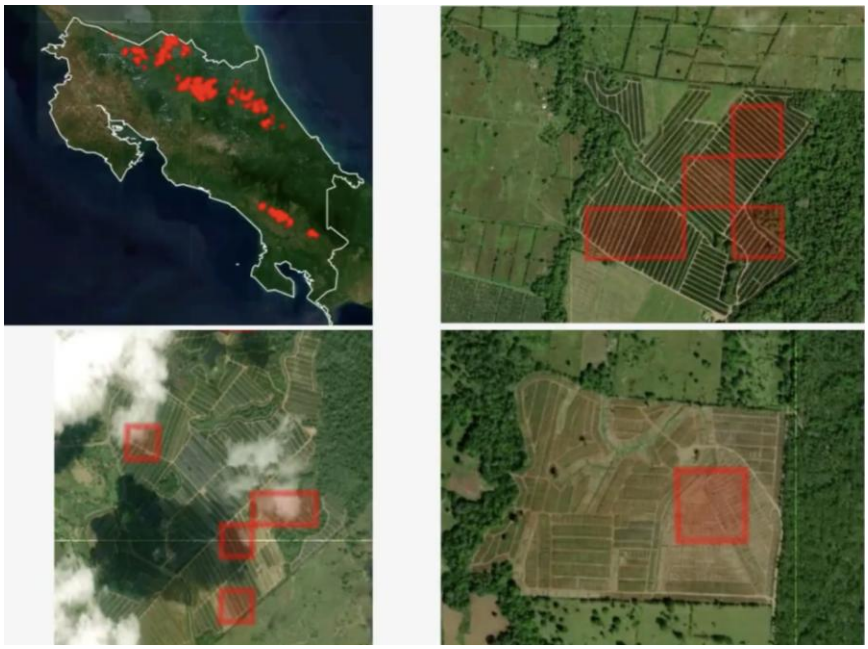
Tracking Land, Food, & Health with AI

Earth Genome is expanding Earth Index, a geospatial AI platform that maps land-use and monitors environmental impacts from many activities, including food production. The project supports scalable, localized model tuning and integrates directly with tools used by researchers and investigators.





Applications

- **Food System Monitoring:** Identifies cattle feeding operations and tracks deforestation from commodity production.
- **Environmental & Health Intelligence:** Correlates land cover change with pollution, disease trends, and conservation priorities.
- **Data-Driven Policy Support:** Helps governments and communities respond to ecosystem and health risks.

Key Features

- **Foundation Models for Land Use AI:** Trained on global geospatial datasets, adaptable to local conditions.
- **Multisource Data Fusion:** Integrates satellite, environmental, and agricultural data in one system.
- **Localized Fine-Tuning:** Enables context-specific analysis across diverse geographies and issues.

Mapping Agriculture and Health Impacts

Detects land-use changes like crop expansion and links them to water quality and public health risks.

Streamlined Scientific Workflows

Offers API access for use in QGIS, Jupyter Notebooks, and other geospatial analysis environments.

Team:
Earth Genome (Ben Strong, Edward Boyda, Hutch Ingold, Mikel Maron)

Technical Challenges

One significant technical challenge the team faced was in data preparation, specifically scaling their Sentinel-2 composite and embeddings pipeline to be globally repeatable, high-quality, and cost-efficient. They completely overhauled their mosaic generation system, introducing a three-pass selection and optimization process using Scene Classification Layers (SCLs) to improve scene selection, eliminate snow/clouds, fill gaps, and reduce no-data pixels. These changes, along with codebase optimizations (e.g., switching to GDAL, refining *NumPy* usage, better chunking in *Dask/Xarray*), reduced compute costs by ~65% and significantly improved imagery quality. The team also migrated workloads to AWS Graviton instances, realizing further cost and performance gains.

Additionally, they encountered the complexity of managing shared, evolving data science notebook environments as collaborative products. This is an underrated challenge in interdisciplinary, open-source research.

Observations and Lessons Learned

The generative AI + geospatial space is vast and rapidly evolving, encompassing diverse technical methods and applications. It may prove valuable to "map the map" and identify where specific GenAI methods intersect with Earth Observation and analysis tasks.

Cross-team engagement reinforced the importance of grounding technical experimentation in real-world impact, ensuring that innovation serves practical, mission driven outcomes.

The team also observed that different foundation model performance varies significantly by task and context, motivating the development of use-case-specific benchmarks, especially for their search and monitoring pipelines.

Next Steps

The Mordecai Lab is now completing searches for pineapple plantations in Costa Rica and incorporating the resulting land-use footprints into dengue transmission models. These findings will be shared with the Costa Rican Ministry of Health and submitted for publication.

Customers/Operations

This work is currently being operationalized by the Mordecai Lab in their disease modeling workflows. Results are being aligned with real-world public health interventions in Costa Rica. Looking forward, Earth Genome aims to apply this AI-powered agricultural monitoring approach to other agricultural practices, regions, and public health challenges. Earth Genome welcomes collaborators at the intersection of agriculture, health, and remote sensing to help expand the scope and impact of their work.