Earth Genome

SOURCE COOPERATIVE

SHARES THROUGH SOURCE

INTRODUCING EARTH GENOME: DATA, SCIENCE AND PEOPLE FOR THE PLANET

Earth Genome focuses on data, science and people for the planet. Earth Genome was started to translate increasingly massive Earth data into useful, accessible, affordable analysis to enable more rational and sustainable use of natural resources. Their expertise is working with AI and Earth data. They build tools to create data, apply science, influence, guide and measure effective actions on the most pressing issues of our times.

CREATING GLOBAL EARTH IMAGERY AND EMBEDDINGS

One of Earth Genome's projects is the Earth Index application, a search tool for the environment, harnessing satellite imagery and machine learning to extract an area's "genetic signature." Earth Index helps to solve one of the greatest challenges for conservationists which is detecting environmental harms before it's too late. Earth Index enables users to easily find and act on damaging activities, like illegal mining or illegal farm expansion, in minutes not months.
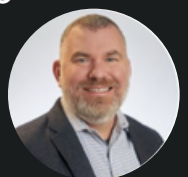
Earth Index uses multiple large-scale datasets and a global, cloud-optimized satellite imagery composite, which is about 60 terabytes, and approximately 3.5 billion embeddings. Their pipelines have been highly optimized, resulting in substantial cost savings which are orders of magnitude cheaper than other methods, so they can responsibly steward resources to rapidly process new data and new models when available. These resources have potential to be widely useful. Since Earth Genome is committed to being a contributor to the Earth+AI community they were driven to openly share this processed data for use in new and unexpected ways.

EARTH GENOME SHARES THROUGH SOURCE COOPERATIVE

Earth Genome believes that foundational planetary data should be open and accessible. That's why they chose to release these datasets on Source, which they see as a platform dedicated to fostering collaboration around open data for public good. Source provides the infrastructure necessary to host and share these large-scale datasets, enabling the community to easily access, explore, and build upon them.

"Source has allowed us to open up all of this data to really anybody in the world... we would have been hesitant to do this without Source. We get people from local communities, indigenous groups, and others from almost every continent, who have reached out to us about this data."

Tom (Hutch) Ingold , Earth Genome, CTO

SHARING THROUGH SOURCE: KEY BENEFITS

- Global Accessibility: Source enables worldwide access to processed datasets that would otherwise be financially or technically out of reach for many organizations, particularly nonprofits and communities in developing regions.

- Simplified Distribution: Source's catalog infrastructure and self-service model eliminated the need for Earth Genome to build and maintain their own hosting infrastructure and web interfaces.

- Community Impact: The open availability has attracted inquiries from journalists, indigenous groups, local communities, and researchers across nearly every continent.

SHARING THROUGH SOURCE: IMPACT

- 60+ terabytes of processed satellite imagery made freely available
- Trillions of pixels processed and shared with the global community
- Global reach with users and inquiries from nearly every continent
- Cost savings allowing Earth Genome to focus resources on core mission activities

LOOKING FORWARD

Earth Genome values Source cooperative model, which aligns with their commitment to open data and community-driven approaches. The organization appreciates being part of a collaborative ecosystem where data publishers work together to democratize access to valuable geospatial resources.

EARTH GENOME DATASETS ON SOURCE

Sentinel 2 Cloud Free Temporal Mosaics
EarthIndex Embeddings from Sentinel 2 Temporal Mosaics
EarthIndex yearly imagery mosaics

FIND OUT MORE ABOUT EARTH GENOME

Learn more about Earth Genome: www.earthgenome.org

ABOUT SOURCE COOPERATIVE

Source Cooperative is a data publishing utility that allows trusted organizations and individuals to share data using standard HTTP methods.

A core belief behind Source is that it should be much easier for people and organizations to share data on the open web. Commoditized cloud object storage has made it easier than ever to store and share data, but it has not made it easier to discover and access that data. Source is an attempt to change this by making it easy for data providers to publish data in an object storage service and make it available on the open web without needing to run their own server, create a data portal, an API, or a dashboard.
Find out more at https://source.coop/