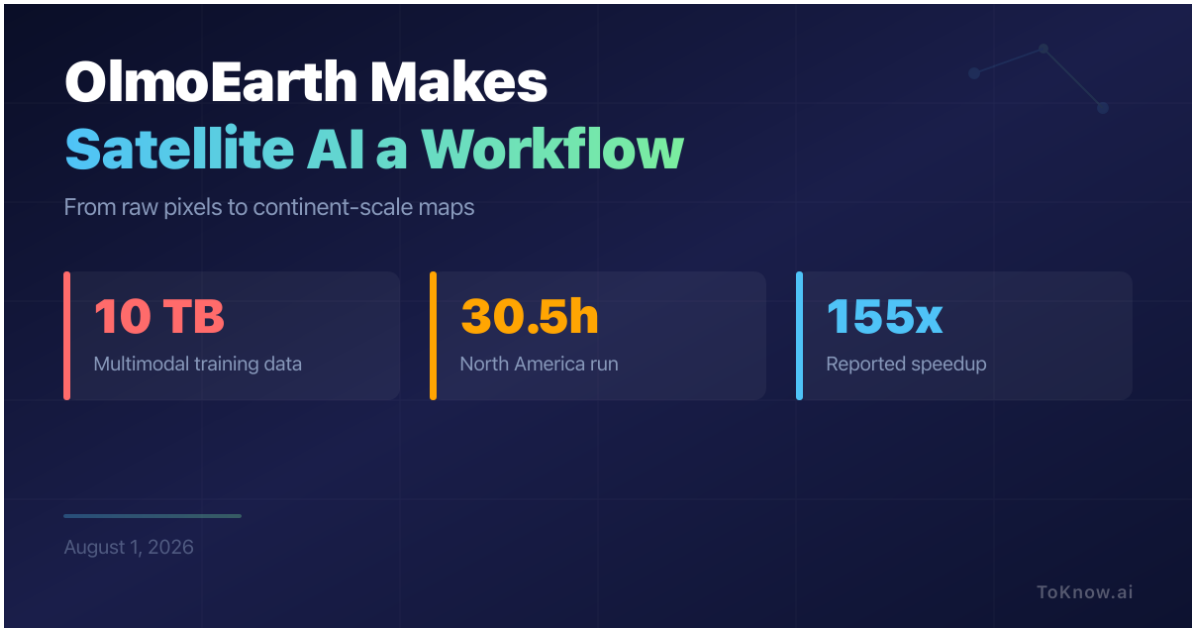
OlmoEarth Makes Satellite AI a Workflow

Kabui, Charles

2026-08-01

Read at ToKnow.ai



OlmoEarth is Ai2's platform for organisations that need satellite maps but do not have a large geospatial engineering team. Its open Earth-observation models are trained on roughly **10 terabytes** of multimodal satellite data, including Sentinel-1, Sentinel-2, Landsat, and derived maps. The hosted Studio handles data import, annotation, fine-tuning, deployment, and export, while the API supports dataset and prediction jobs (OlmoEarth; OlmoEarth documentation). The difficult part is not only running the model. The pipeline must find

usable scenes, reproject them, read only the needed windows, and stitch the predictions into a map.

Ai2 reports a North America wildfire-risk run using **19,600 CPUs**, **994 GPUs**, and more than **168 GB/s** of network throughput. It reduced an estimated 4,737 serial compute hours to **30.5 hours of wall-clock time**, a reported **155x speedup** (Ai2's engineering account). That is parallel elapsed time, not a 155x cheaper or faster GPU, and the cost still depends on resolution, model size, cloud prices, and data transfer. Google Earth Engine and Microsoft's Planetary Computer already provide large geospatial catalogues and cloud APIs (Earth Engine; Planetary Computer). OlmoEarth's difference is the model-specific workflow. Its open code and dataset support self-hosting, but that path leaves preprocessing, monitoring, licences, and infrastructure to the user (code; dataset).

Read More: Four AI Projects Put Structure Before Scale

Sources:

- The OlmoEarth Platform
- OlmoEarth
- OlmoEarth documentation
- OlmoEarth pretraining repository
- OlmoEarth pretraining dataset
