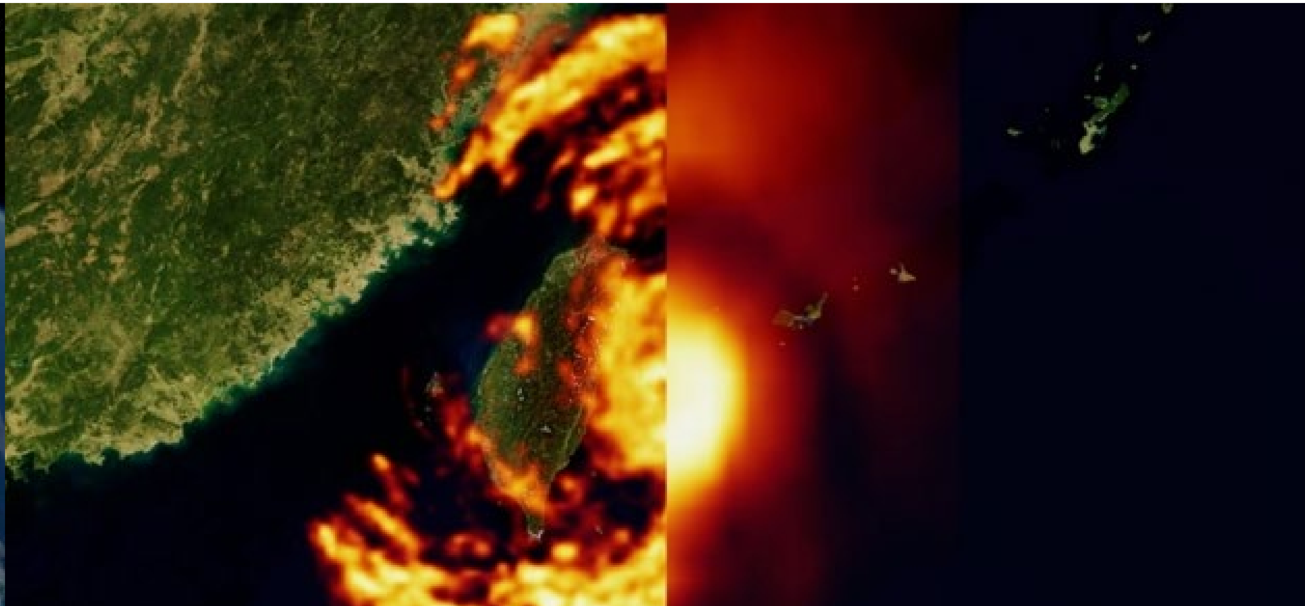
# Deep Learning and Foundation Models for Earth and Planetary Sciences

Jichao Fang

Department of Earth, Atmosphere and Environment

4/10/2026

# Earth Observation - Why do we care?

- Earth observation and remote sensing are integral part of everyday life
- We see it everywhere, use it everywhere
- For computer science, it archives petabyte-scale imagery data with rich spatial, temporal, and spectral structure --- a natural playground for representation learning, self-supervised methods, retrieval, and distribution shift problems.

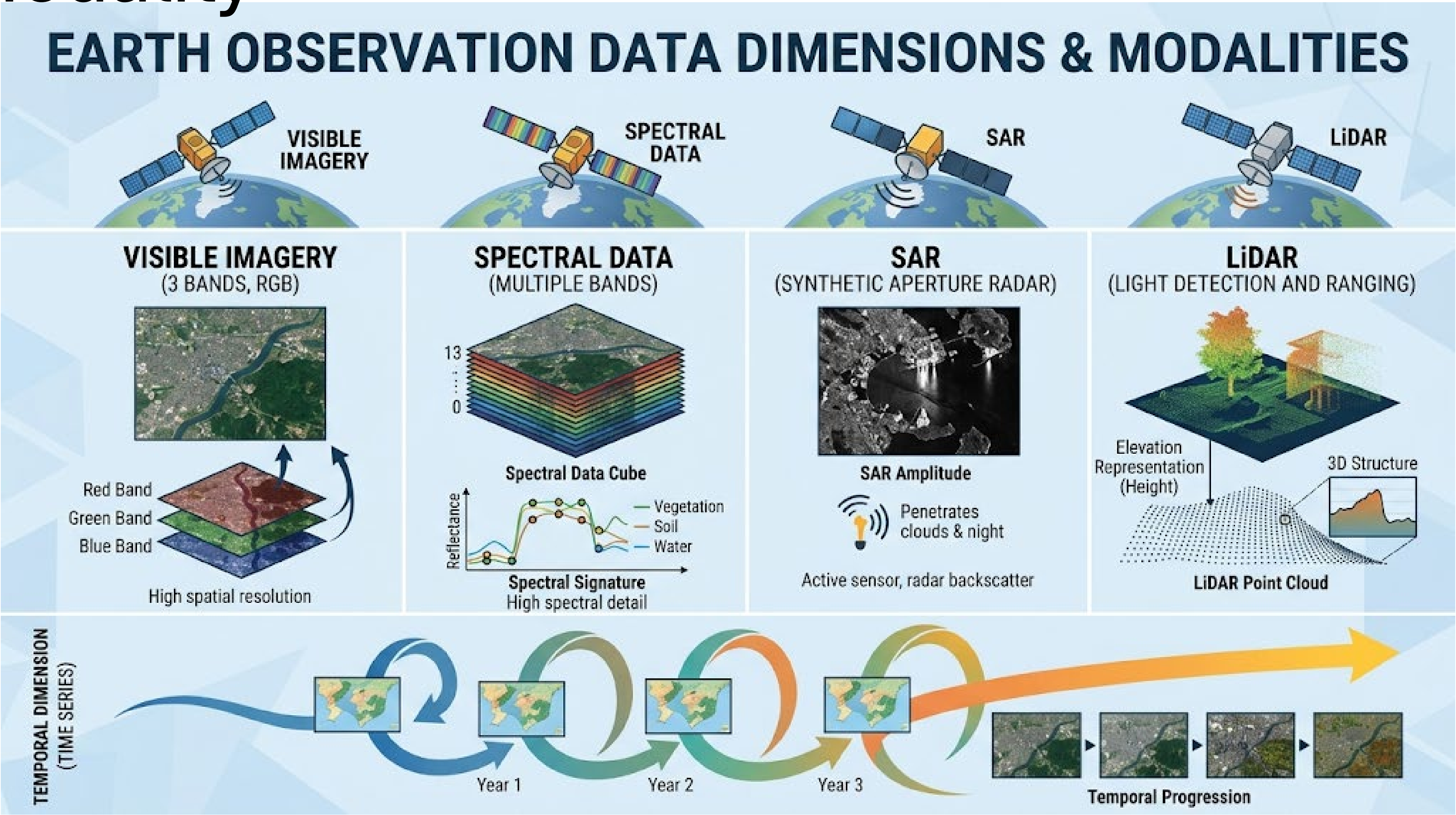
# The tasks

- Object (vehicles, buildings) detection
- Semantic segmentation (landcover, e.g., segment forest, crop field, water body)
- Hazard mapping (wildfire, flood, landslide)
- Temporal: Change detection (movement, scene changes, monitoring)
- Regression and prediction (crop yield, soil moisture)

# Data Modality

- Visible (i.e., RGB imagery)
- Spectral data (more than RGB)
- SAR
- LiDAR

- One more important dimension: Time




Generated by NanoBanana

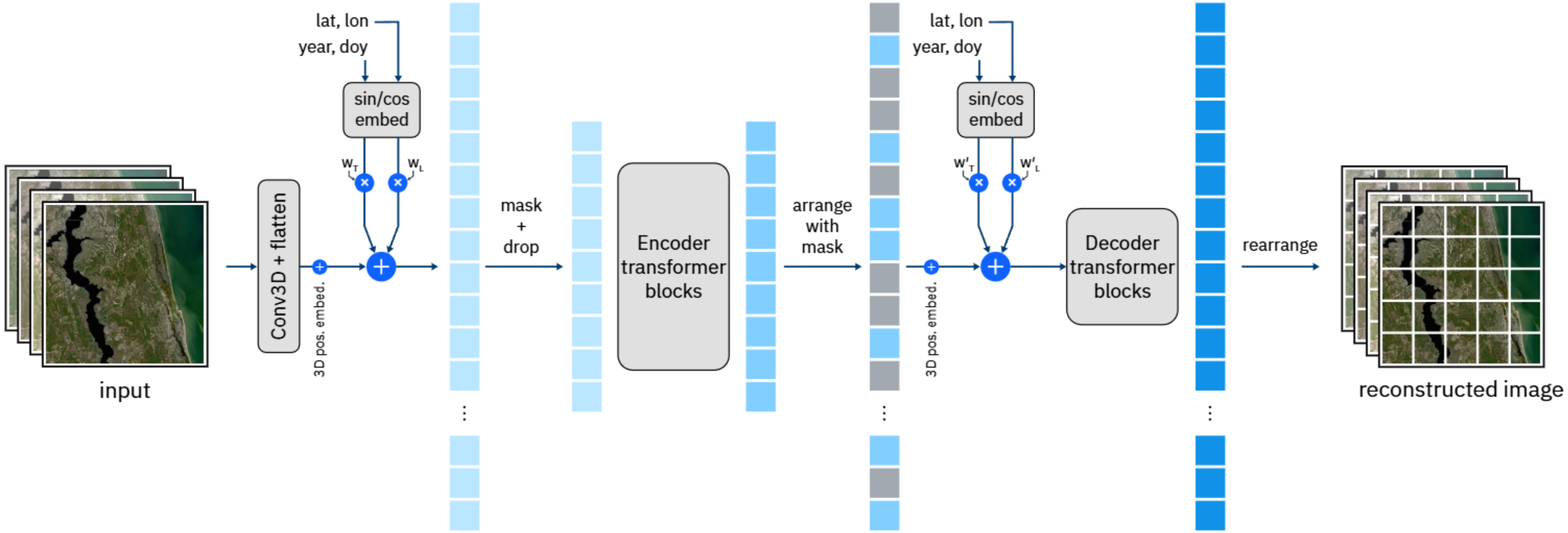
# Data Volume

- NASA's EOSDIS archive surpassed 123 **petabytes** at the end of 2024
- Planet Labs' constellation captures **~30 TB of imagery per day**
- Globally, total EO data collection is estimated at **~807 PB** and growing by at least **~100 PB per year**
- None of it comes with labels

# Geo-Foundation Models

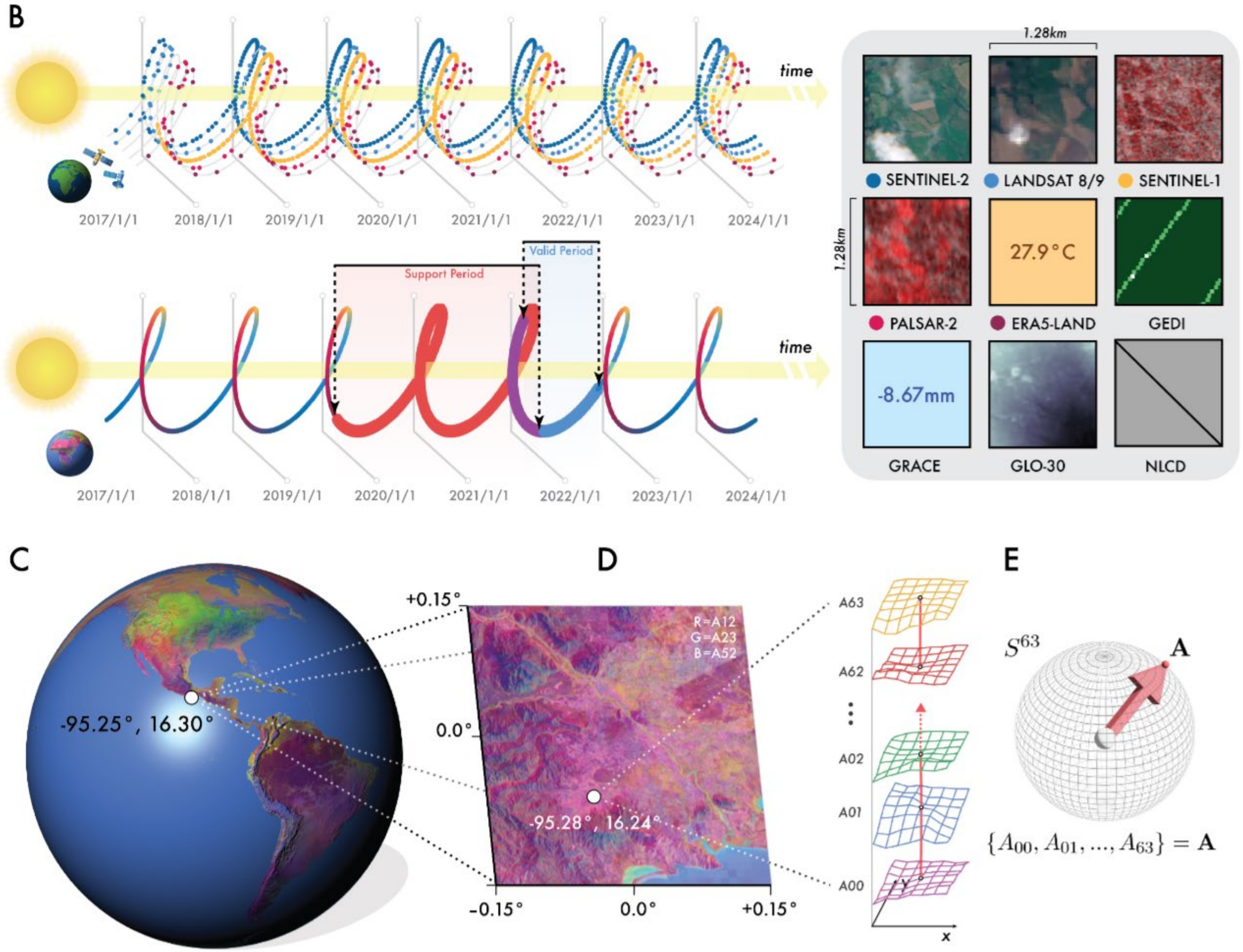
- Foundation models
  - NLP: all modern llms
  - CV: from ImageNet-pretrained CNN (ResNet, VGG, etc) to ViTs (DINOs, CLIP)
  - World models
- Geo-foundation model currently lean towards **vision** foundation models, but mostly explicitly incorporating spectral, temporal, and geographic(location) information

# Ex1. Prithivi (NASA/IBM)

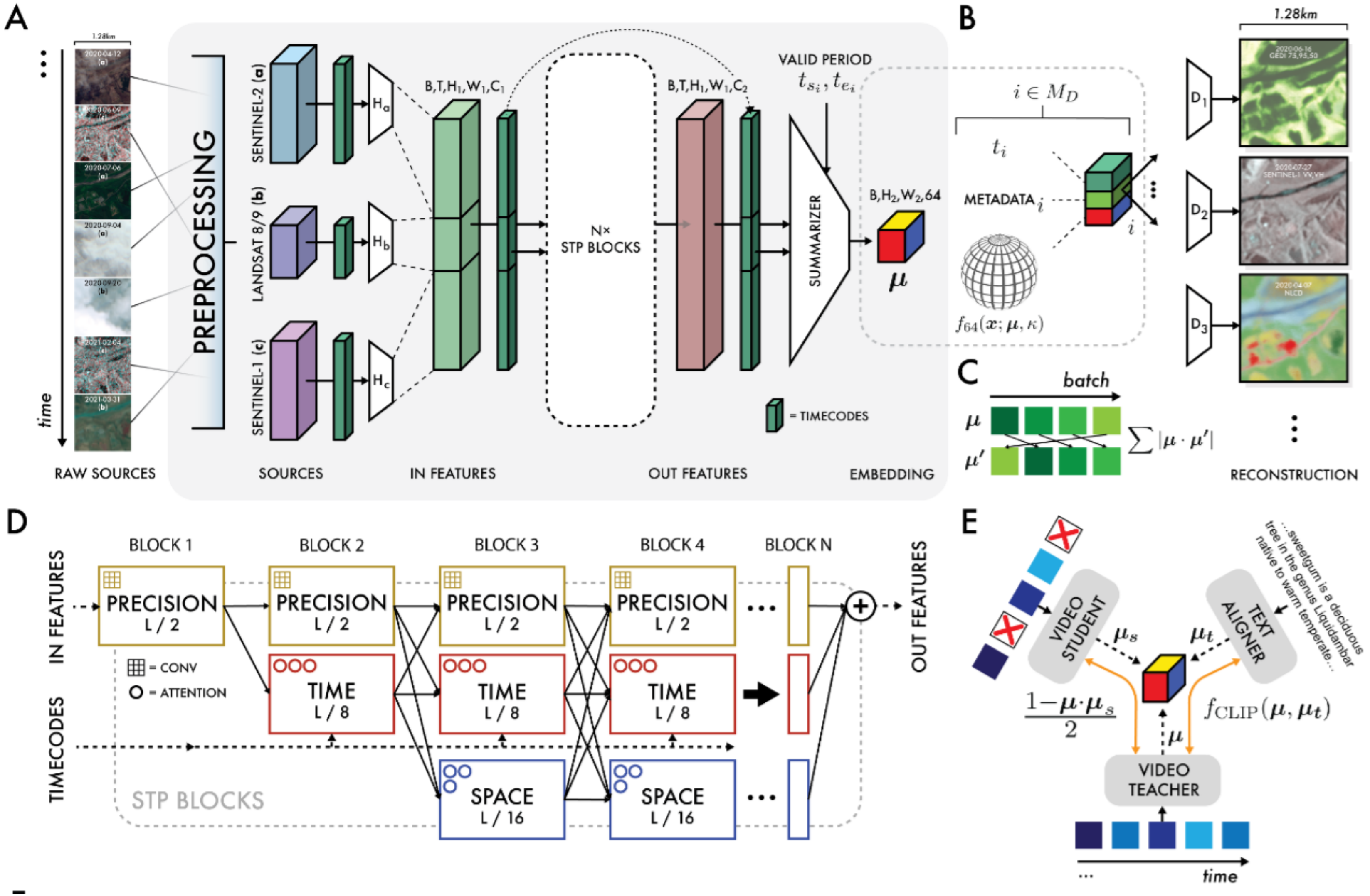



Core: Masked AutoEncoder (MAE)

# Ex2. AlphaEarth (Google DeepMind)



# Ex2. AlphaEarth (Google DeepMind)



# Ex2. AlphaEarth (Google DeepMind)

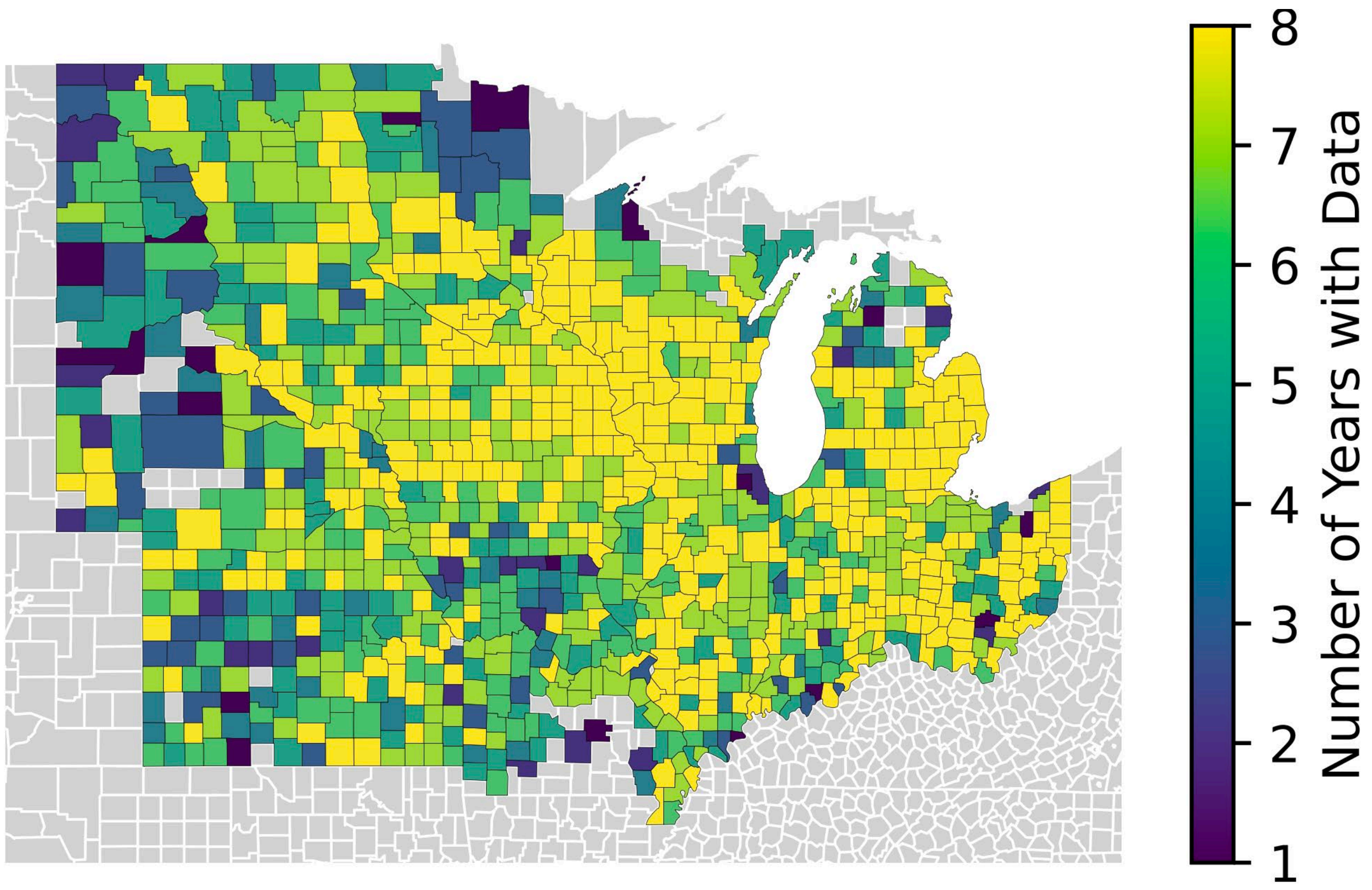
- Most transformative because it jumps out of conventional way of a single model family that requires scientist to run ad-hoc inference
- It releases embedding as analysis-ready data product at high spatial resolution (10m)
- Highly compact (64-D vs ViT-Base 768-D)
- “A representation of Earth’s land surface at 10-meter scale”

AlphaEarth Foundations helps map our planet in unprecedented detail — Google DeepMind

# Ex3. AEF for Crop Yield Estimation

**Data Sources**

- Features: AEF 64-D annual embeddings via GEE
- Crop masks: USDA CDL (corn = class 1, soybean = class 5)
- Labels: USDA NASS Quick Stats (survey-based, end-of-season yields)
- Aggregation: county-level mean of crop-masked embeddings
- **64 predictors per county-year record**




- **992** Counties
- **6,350** County-Year Records
- **2017–2024** Temporal Coverage

# Ex3. AEF for Crop Yield Estimation

### Leave-One-Year-Out (LOYO)

- Hold out one calendar year → test on it
- Train on remaining 7 years
- Prevents within-year info leakage
- Unbiased under out-of-year covariate shift
- **Prediction ≠ forecasting (not chronological)**
- Also report roll-forward (train ≤t−1, test t) for 2019–2024

### Baselines

- County-mean: training-year avg yield per county
- NDVI+PRISM+GDD/CDD: classical phenology stack with MODIS NDVI, precip, growing/cooling degree days
- Model sweep: Linear, Ridge, Lasso, ElasticNet, RF, GBT, MLP, KNN, SVR
- **Metrics: $R^2$, RMSE (bu/ac) per year**

# Ex3. AEF for Crop Yield Estimation

| **Model** | **Mean Test $R^2$** |
|---|---|
| LinearRegression | 0.6968 |
| Ridge | 0.6980 |
| Lasso | 0.6974 |
| ElasticNet | 0.7156 |
| SVR_rbf | 0.8077 |
| LinearSVR | 0.7223 |
| RandomForest | 0.7695 |
| GradientBoosting | 0.7546 |
| MLPRegressor | 0.6761 |
| KNeighborsRegressor | 0.7140 |

TABLE I

LEAVE-ONE-YEAR-OUT TEST (CORN): MEAN TEST $R^2$ BY MODEL, THE BENCHMARK USE *UNTUNED* MODEL SO THE BEST RESULT IS DIFFERENT THAN THE FINAL ONE WE USED.

**Key Takeaway**

SVR-RBF wins at 0.808 (untuned). Smooth nonlinear effects matter.

**After tuning: $R^2 = 0.825$**

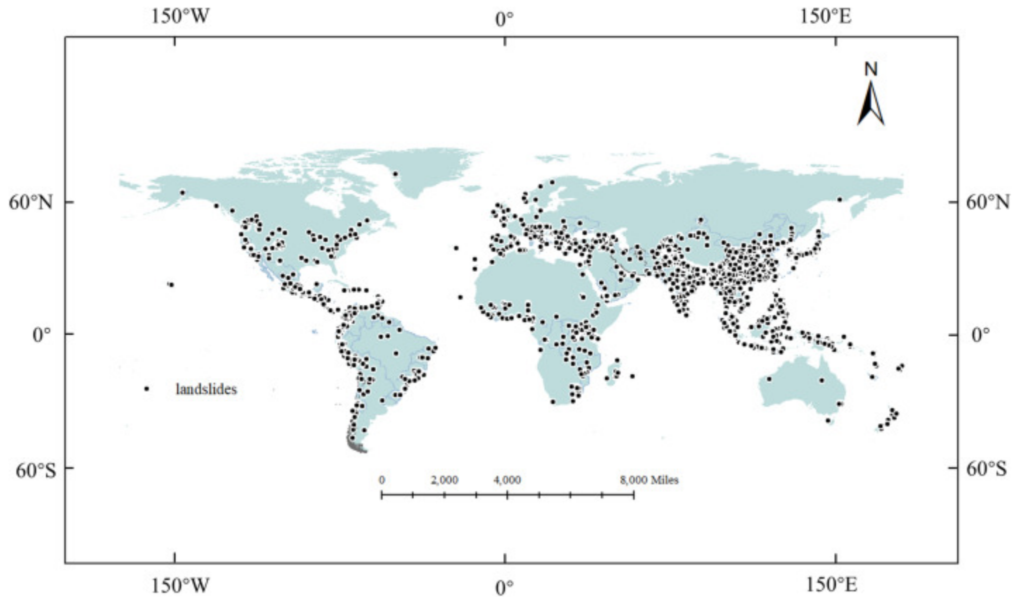
$C \approx 52$, $\gamma \approx 0.007$, $\varepsilon \approx 0.99$

# Ex3. AEF for Crop Yield Estimation

- A simple regression model on AEF embedding achieves very plausible crop estimation result
- AEF embedding are highly linear separatable, making it analysis-ready, and efficient to use

# Ex4. AEF for Landslide Mapping/Prediction

- Landslide is one of the most widespread and destructive natural disasters
  - occurring most frequently in mountainous regions;
  - Many factors contribute to landslides
    - Preparatory factors: lithology, slope, aspect, …
    - Triggering: rainfall, earthquake, human acitivity
  - thousands of deaths & billions $ of damage worldwide annually (Froude & Petley, 2018)

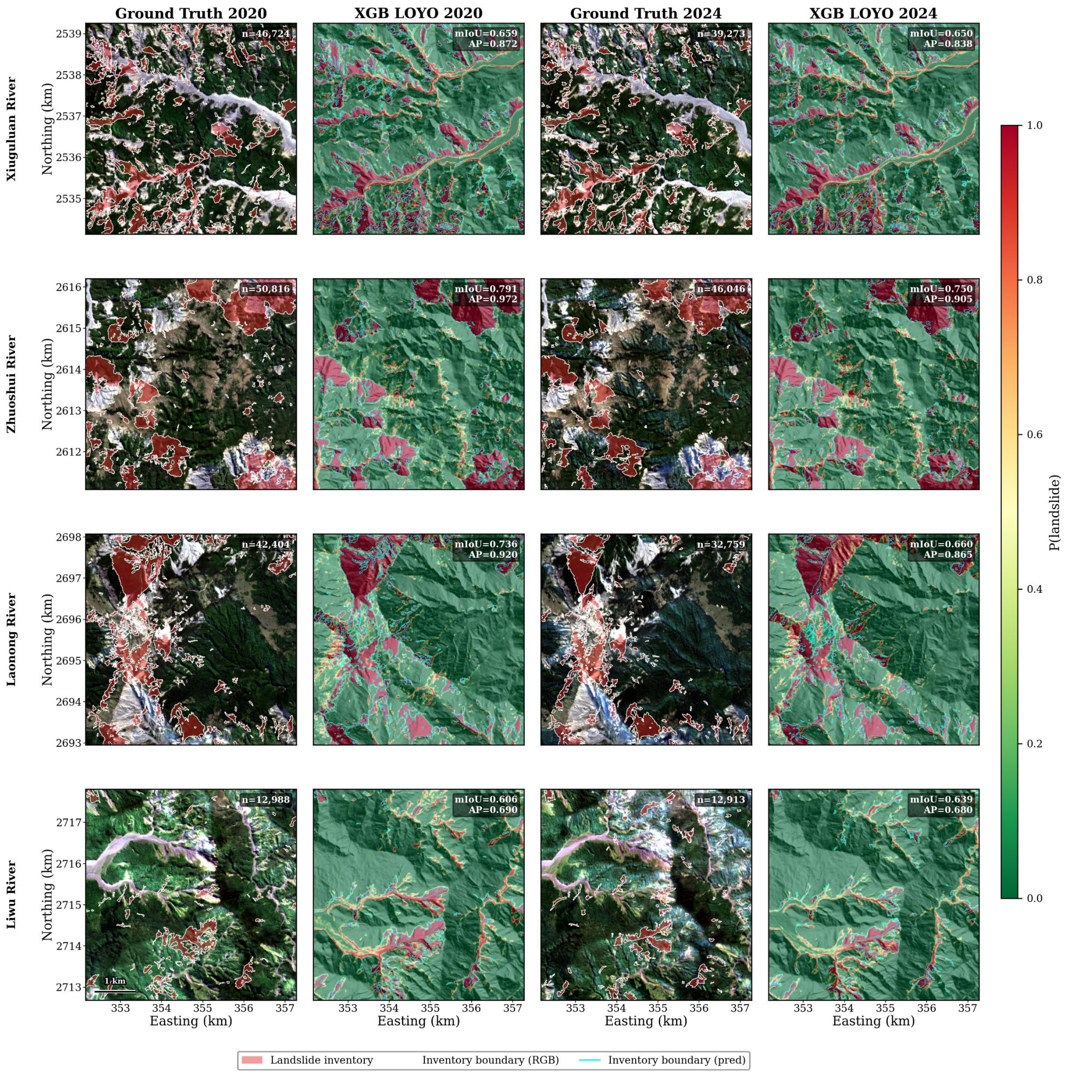



Distribution of the NASA global fatal landslide database (https://svs.gsfc.nasa.gov/4710)

# Ex4. AEF for Landslide Mapping/Prediction

- Extract AlphaEarth embeddings for each grid cell
- Combine embeddings with landslide inventory data
- Train machine learning models to predict landslide susceptibility
- Produce spatial probability maps

# Ex4. AEF for Landslide Mapping/Prediction



# Ex4. AEF for Landslide Mapping/Prediction

- A simple bi-classification model on AEF embedding achieves very high accuracy in landslide indentification
- AEF embedding are highly linear separatable, making it analysis-ready, and efficient to use

- Embed the earth: Highly compressed representation of the Earth, reusable, lightweight pipeline, high scientific value.

# GeoAI – Beyond Vision Models

- AI Weather Forecasting (Surrogates for PDE Solvers)
- Geometric Deep Learning (Earth as a Graph)
- Agentic GeoAI (The Autonomous GIS Analyst)
- Onboard/on-satellite computing
- And to other planets (e.g., Mars)

# Ex5. Agentic AI (trends)

- Traditional GIS: human crafts query → runs tool → inspects → iterates. 1 analyst, 1 task at a time
- LLM agent loop: natural language task → decompose → call GEE/QGIS/Python tools → verify output → report
- Real example: 'Find all agricultural fields that switched to fallow after the 2023 drought in Henan province' — agent writes the GEE script, runs it, maps the result

Article

**Large Language Model Agent with VGI Data for Mapping**

Jiayu SONG, Yifan ZHANG, Zhiyun WANG, Wenhao YU

Research Article

**An autonomous GIS agent framework for geospatial data retrieval**

Huan Ning, Zhenlong Li, Temitope Akinboyewa & M. Naser Lessani

Article: 2458688 | Received 28 Aug 2024, Accepted 21 Jan 2025, Published online: 09 Feb 2025

Cite this article https://doi.org/10.1080/17538947.2025.2458688 Check for updates

*[Submitted on 5 Nov 2024 (v1), last revised 22 Nov 2024 (this version, v4)]*

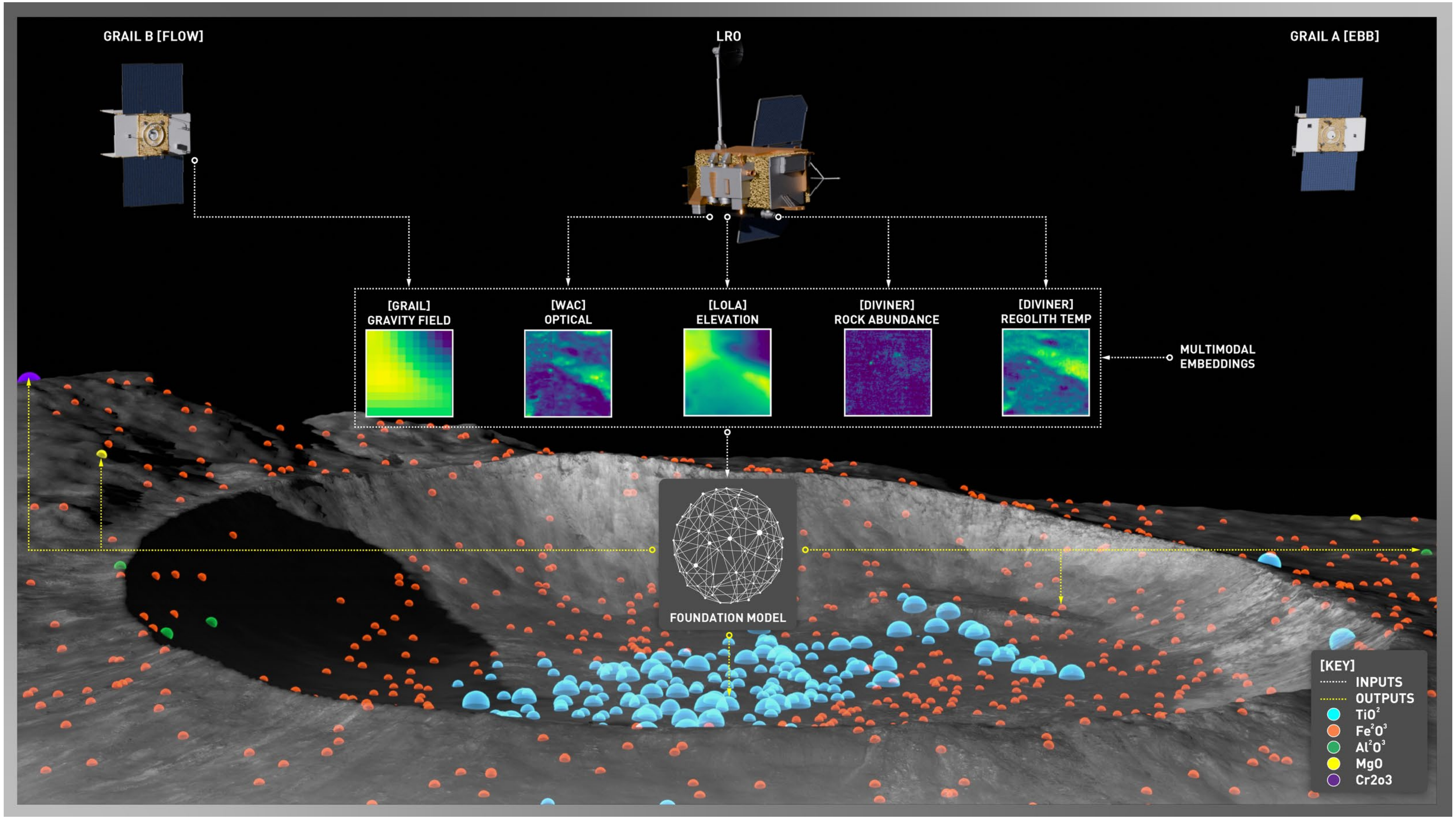
**GIS Copilot: Towards an Autonomous GIS Agent for Spatial Analysis**

Temitope Akinboyewa, Zhenlong Li, Huan Ning, M. Naser Lessani

# Move beyond Earth

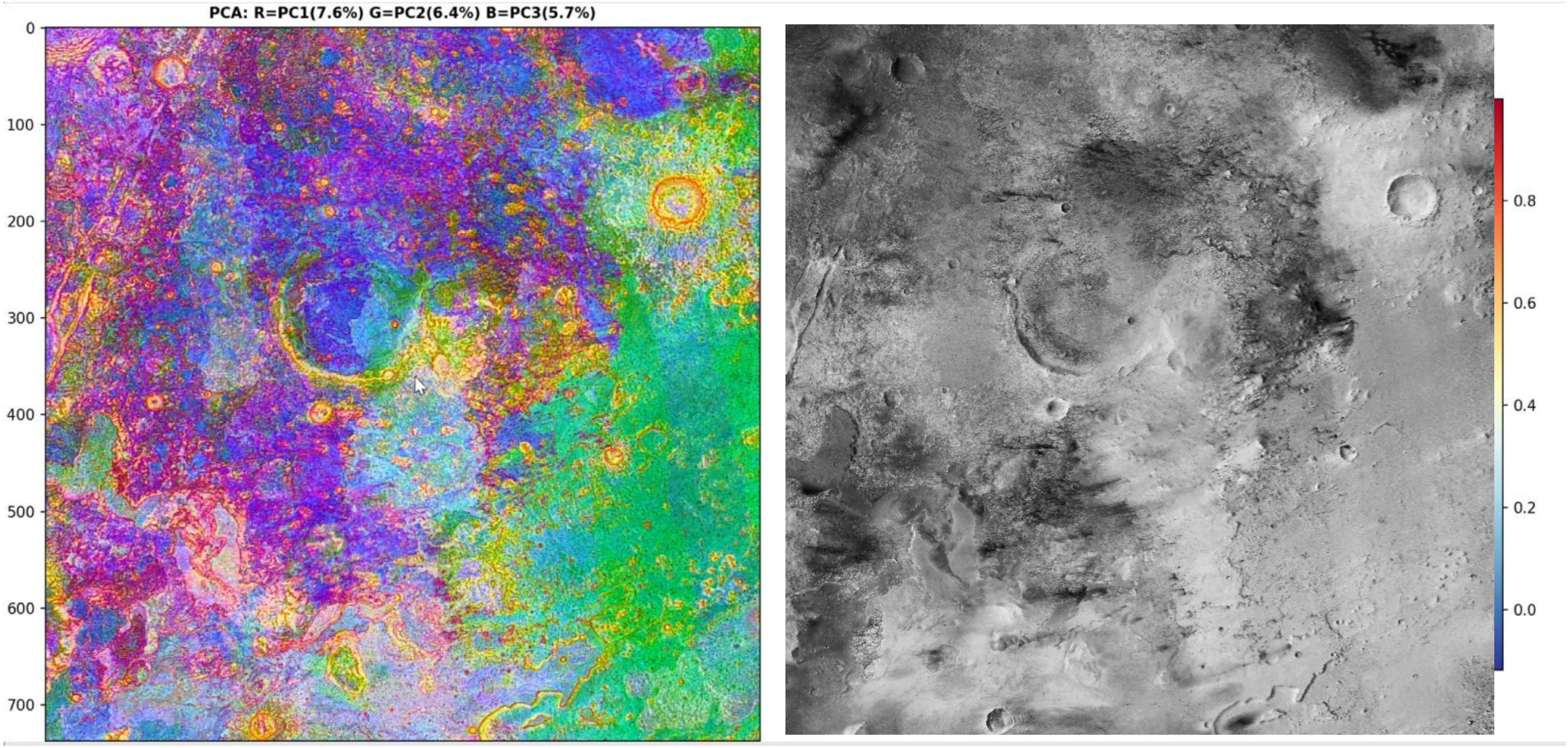
- Space science needs AI even more
- Limited computational capacity, very high latency (autonomous discovery)
- Large amount of imagery, very few can be labeled/verified

# Ex6. A Moon Foundation Model (Luxembourg Space Agency)



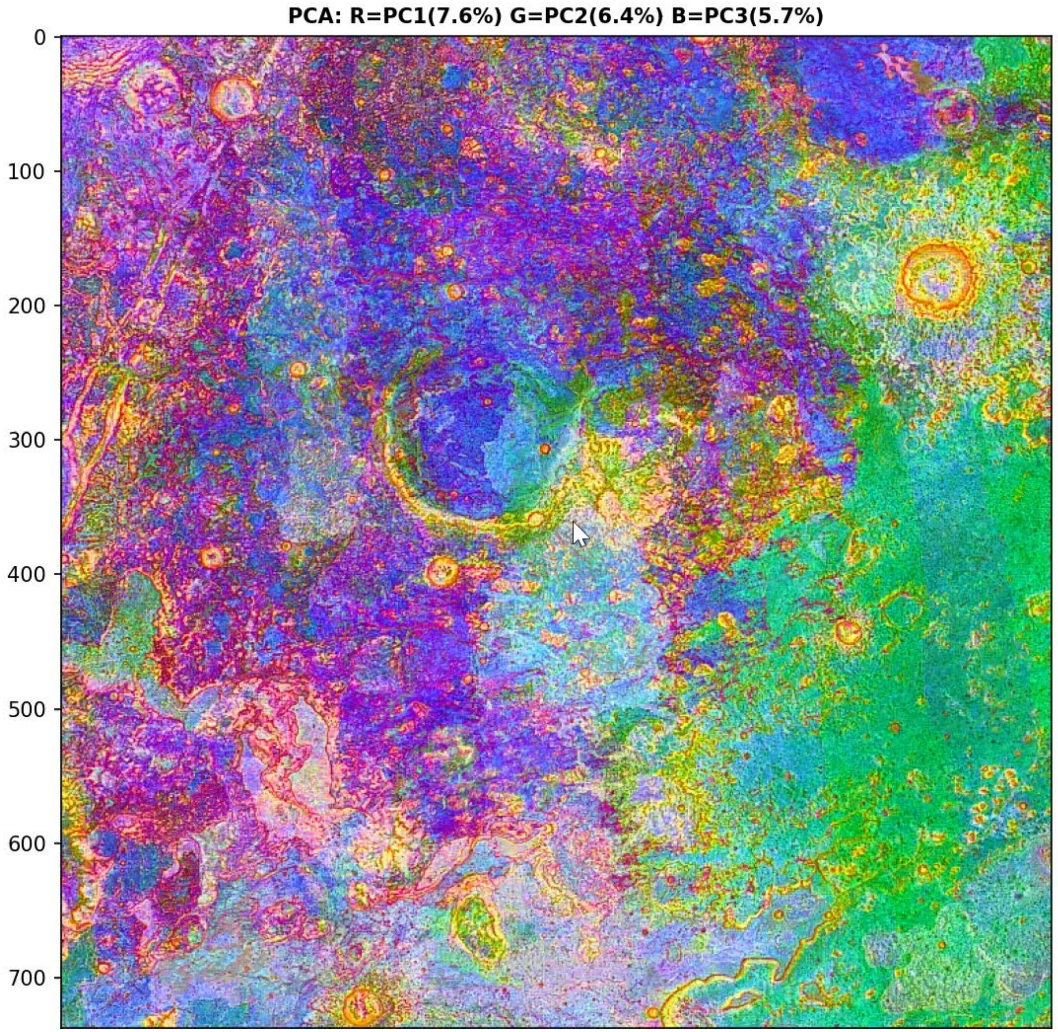
# Ex7. An Global Mars Image retrieval system

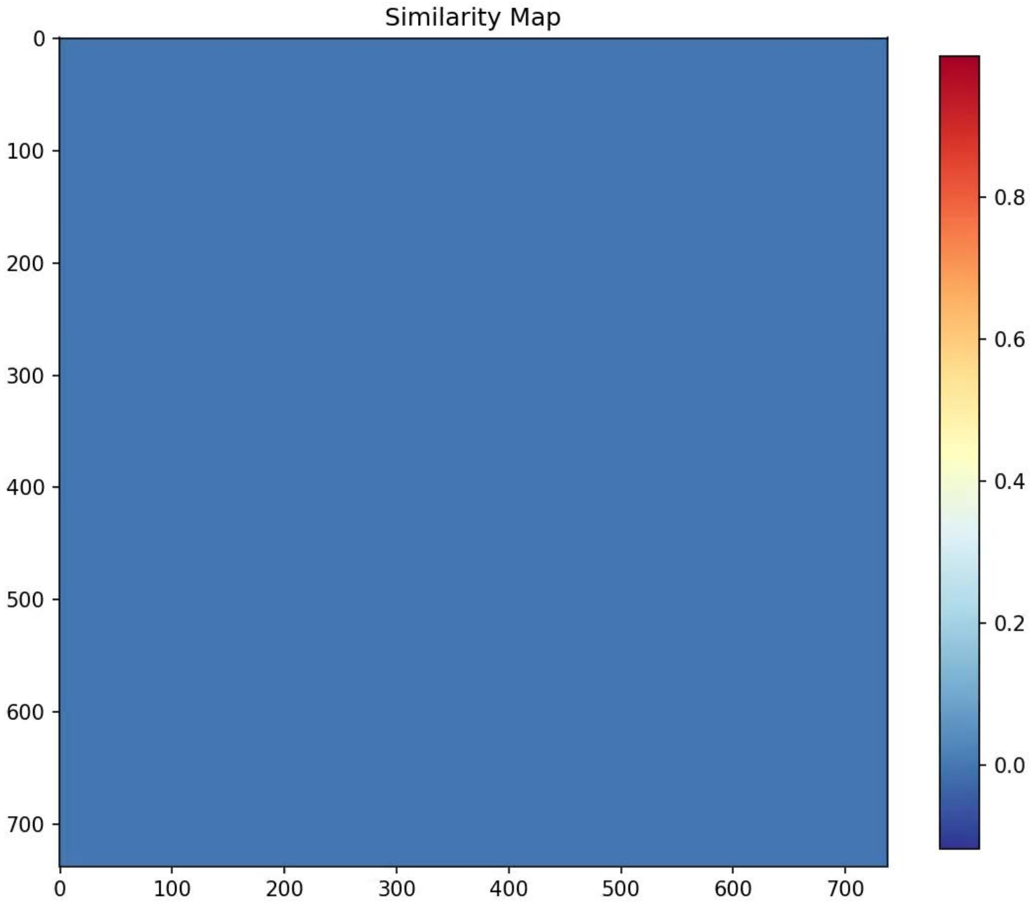
## What does **similarity** look like?



# Ex7. An Global Mars Image retrieval system

## What does **similarity** look like?



# Ex7. An Global Mars Image retrieval system

## **Retrieval System** as Science-Enabling Tool

- Search Large archive of similar images via image embeddings
- Can be used for:
- Localization
- Analog search
- Answers:
- **Where is this crater?**
- What's the distribution of this landform?
- What's the similarity of the two features?

# Ex7. An Global Mars Image retrieval system

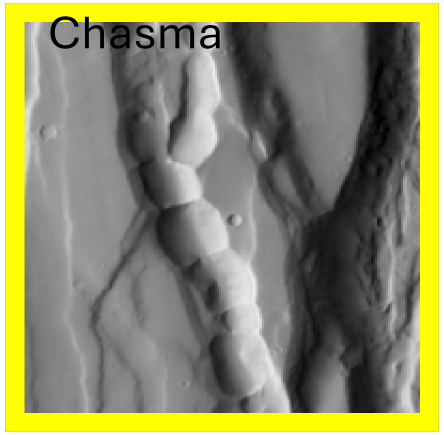
## Localization/Identity Search

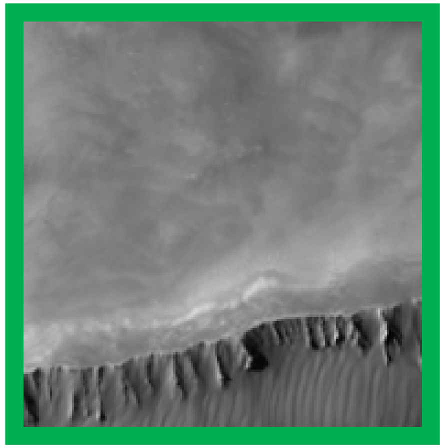
**(Where is this crater?)**

# Ex7. An Global Mars Image retrieval system

## Landform Distribution (What's the distribution of this landform?)

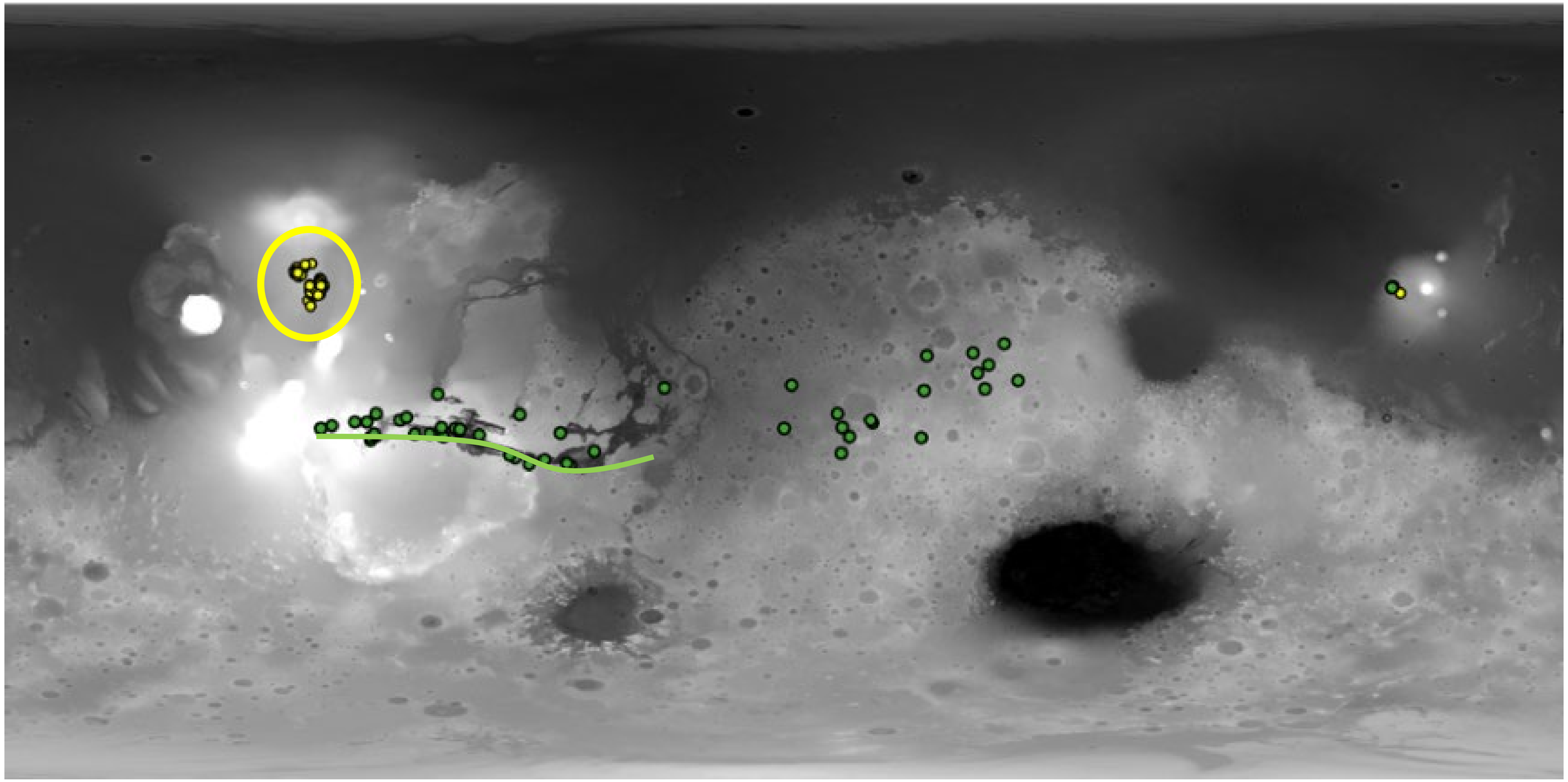
Elysium Chasma





Valles Marineris

Top 50 similarity search results

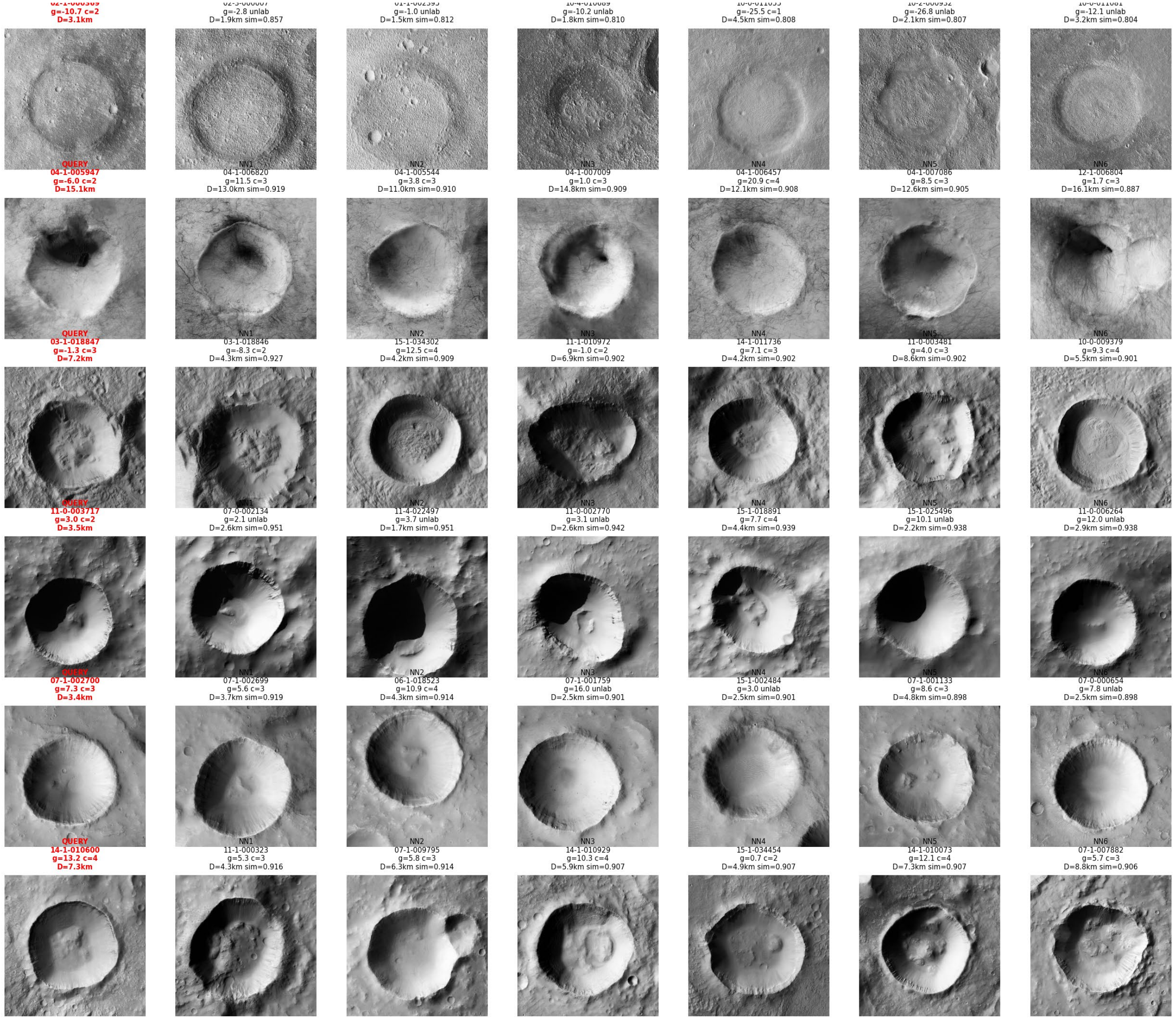


# Search in Action

# Ex7. An Global Mars Image retrieval system

## Similarity Ranking

Providing one example image, searching craters in similar degradation state



# Ex7. An Global Mars Image retrieval system

## Technical Details

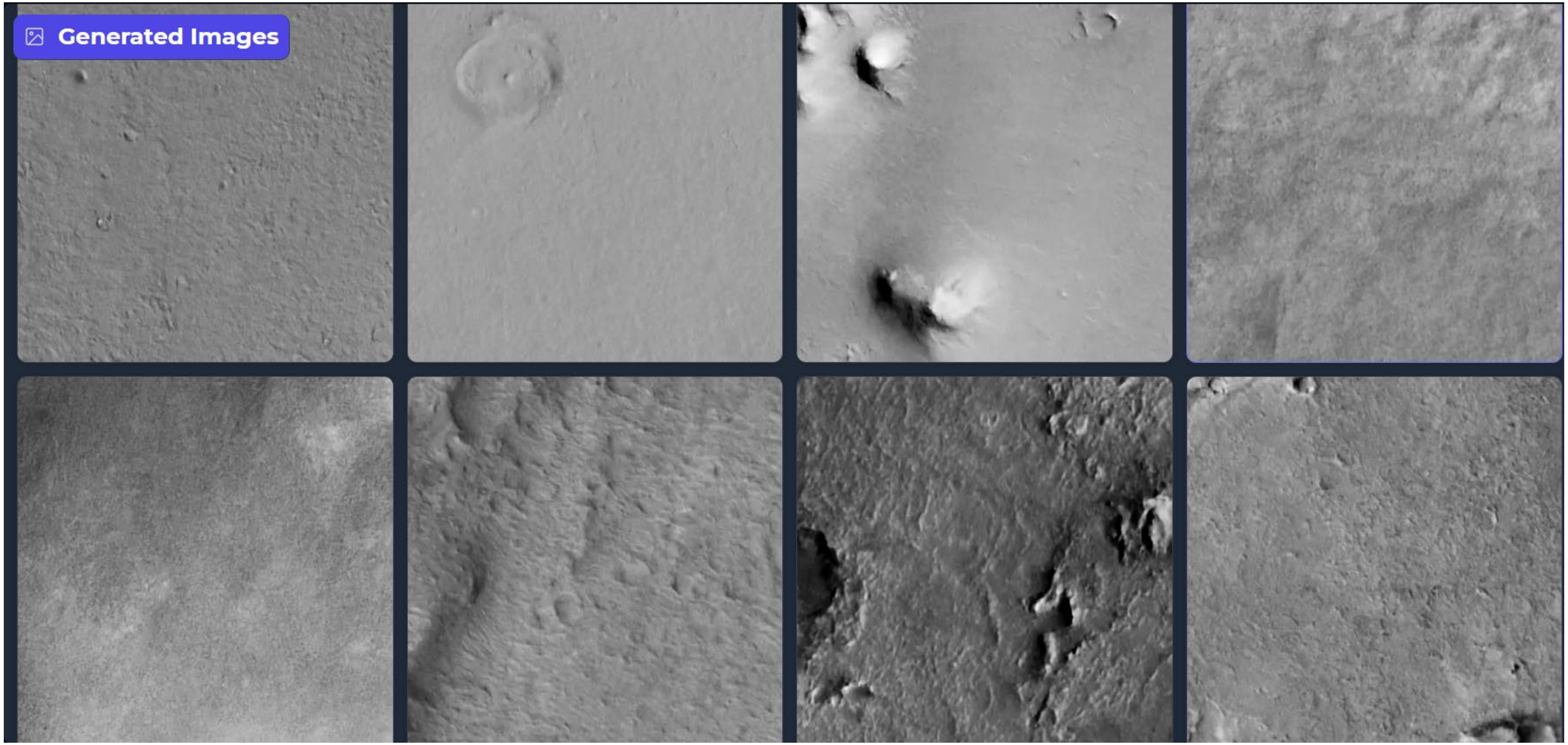
- We indexed >26 million images from Murray CTX (almost exhaustive)
- Each image is passed to Vision Transformer (ViT-Base) model, resulting one pooled 768-D (token) vector for primary indexing
- Additional 32 tokens are used for second stage re-ranking after short-listing
- A single self-manage cloud server handles retrieval elegantly with only 96GB RAM, 16 vCPU cores and 1TB storage. No GPU needed.

# Ex7. An Global Mars Image retrieval system

## Technical Details

- Model: Trained on ~7 million images, MAE ViT strategy
- METIS, 16 nodes training, 1-2 days for 100 epochs.
  - Torchrun handles multimode training nicely with minimum code change requirements

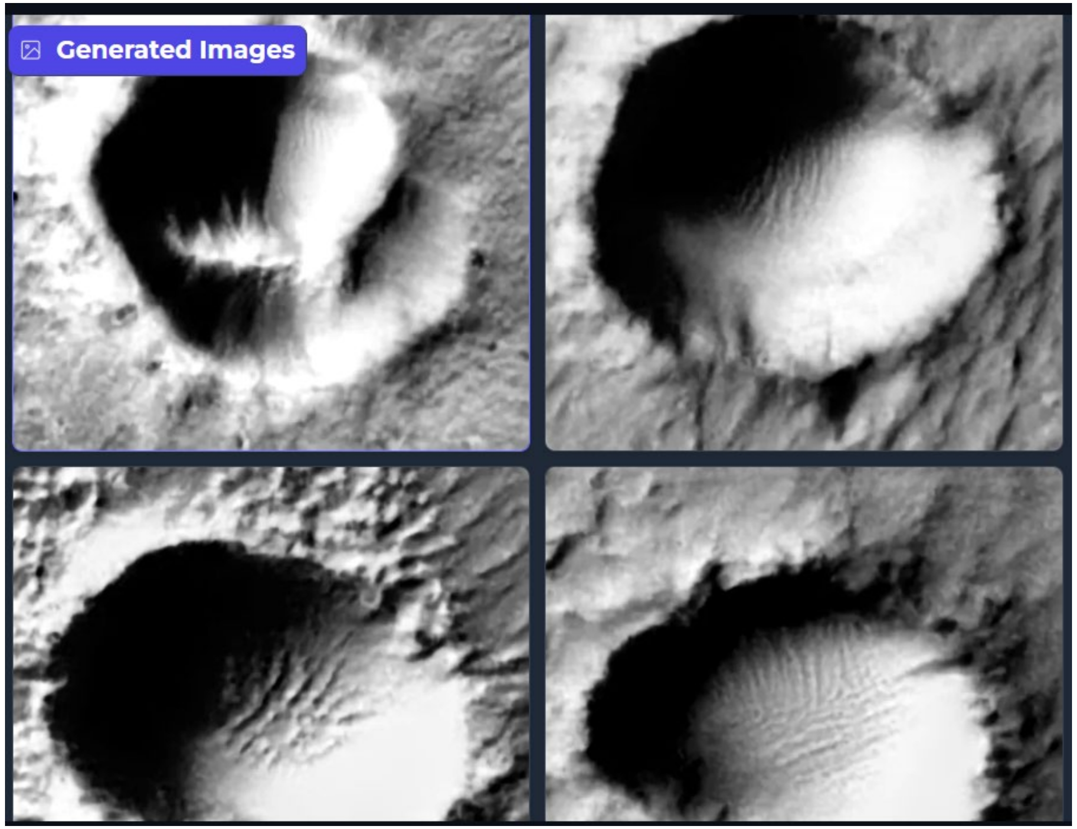
# Ex8. VAE and Diffusion Models for Mars



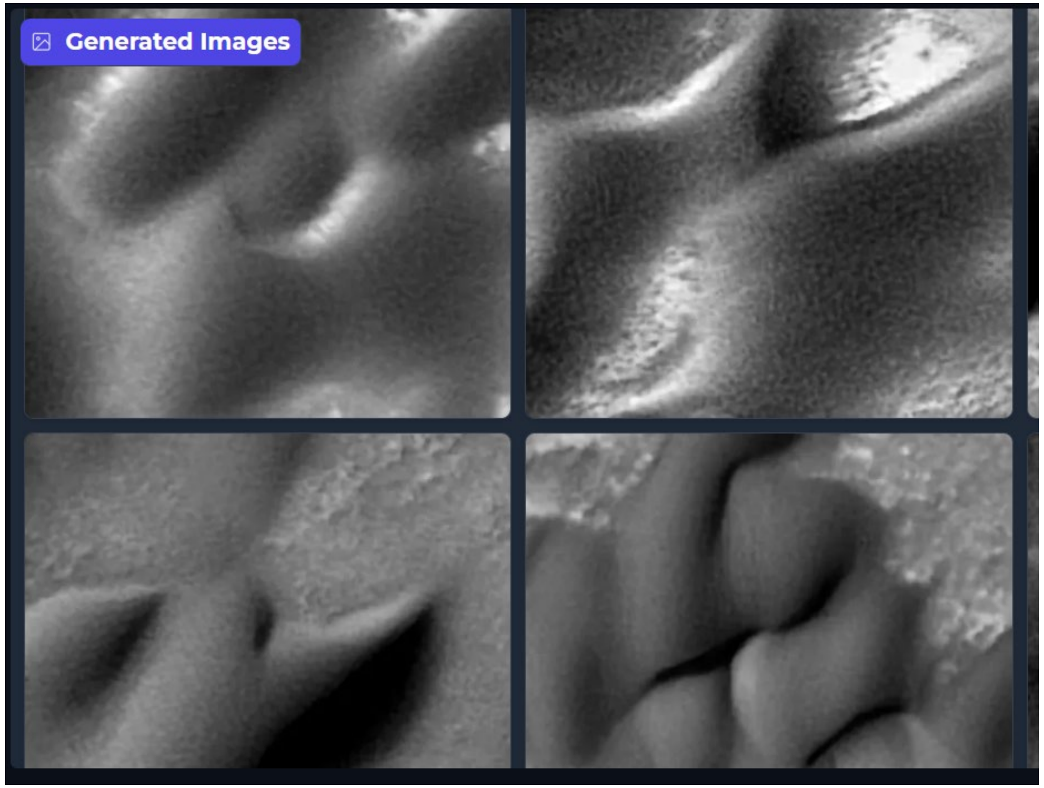
# Ex8. VAE and Diffusion Models for Mars

## Classifier-Free Guidance (CFG)

**Prompt Crater**




**Prompt Dunes**



# Ex8. VAE and Diffusion Models for Mars

**inpainting**

# Thank you!