EO Foundation Models

Why and How to use them

About Tisham

- Senior Engineer at CSIRO Space and Astronomy
- Wrangler of Kubernetes clusters and scientists using them
- Open-source maintainer, Startup CTO, Public servant, Family person, Mangaka

Presentation Outline

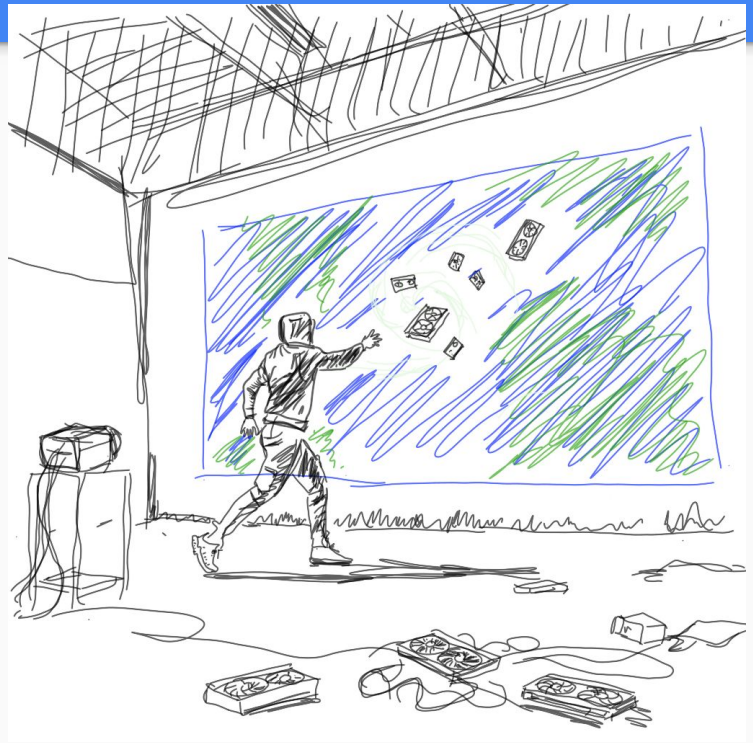
- GeoFM Quick Primer
- Practical Usage
- Making your own GeoFM (why and who should do it ?)
- Fine tuning a GeoFM (when you are not FAANG or NASA/ESA)
- Getting the compute to use a GeoFM
- Running GeoFM’s on Satellites (Distillation)
- Keeping GeoFM’s updated and sustainability
- Future work and Opportunities

What are EO Foundation Models

Geo-Foundation Models (GeoFMs) are a subset of foundation models that focus on explicit representations of spatial primitives such as spatial interaction, spatial stationarity, spatial heterogeneity, and so forth, and encode rich information about places and regions (Janowicz et al. 2020, Li et al. 2021, Mai et al. 2022b).

Image Credit : Firmus Technologies

Just throwing GPU’s at Satellite Imagery



What are GeoFM’s really ?

I see them as a combination of multi-sensor data fusion and analysis ready pre-processing. If someone asked you to create a band reduced ARD from 10 different satellites , a GeoFM is what you would most likely get.

Obligatory Masked Autoencoder (MAE) Slide

Training of the models in presence of limited labelled data takes advantage of the detecting inherent patterns by masking out some of the data

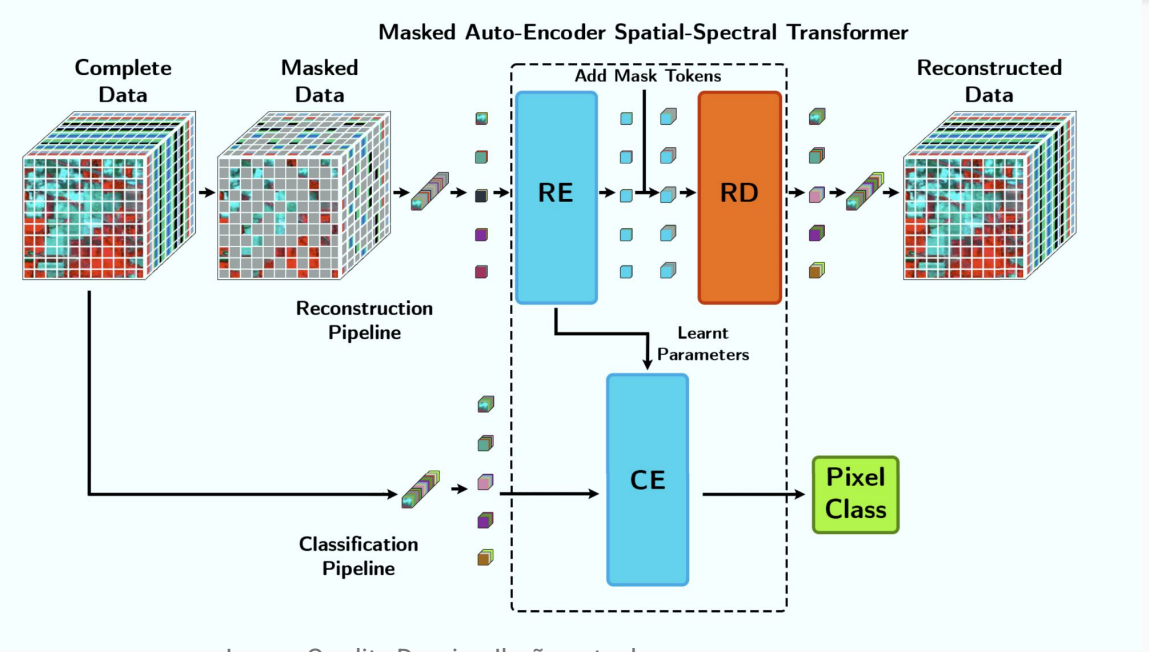


Image Credit : Damian Ibañez et. al.

Obligatory Global Training Datasets Slide

- Foundation models depend on collections of data curated by major space agencies
- The proliferation of Analysis Ready Cloud Optimized (ARCO) satellite imagery is driving forward maturity of the models
- If your data is not ARCO

  You will be LEFT OUT

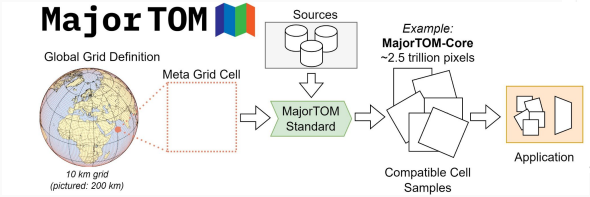
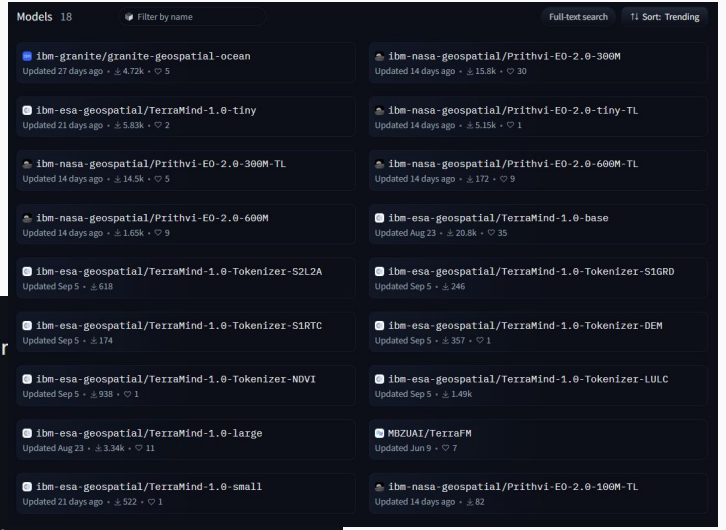


Image Credit : European Space Agency (ESA)

Where to get an Earth Observation FM

- Hugging Face and download weights
- Install from Github
- Install pre-packaged ones from TerraTorch
- Train one from first principles and a given architecture if you hav
  to some HPC



Why are there so many ?

- Like LLM’s there are many GeoFM’s we are the divergent phase where no clear winner has emerged. These are technology push solutions looking for concrete problems.
- The different GeoFM’s solve for different modalities, have slightly different architectures, different training data and sponsoring organizations.

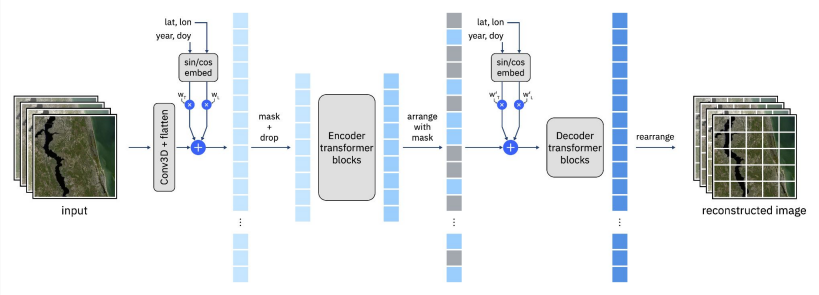


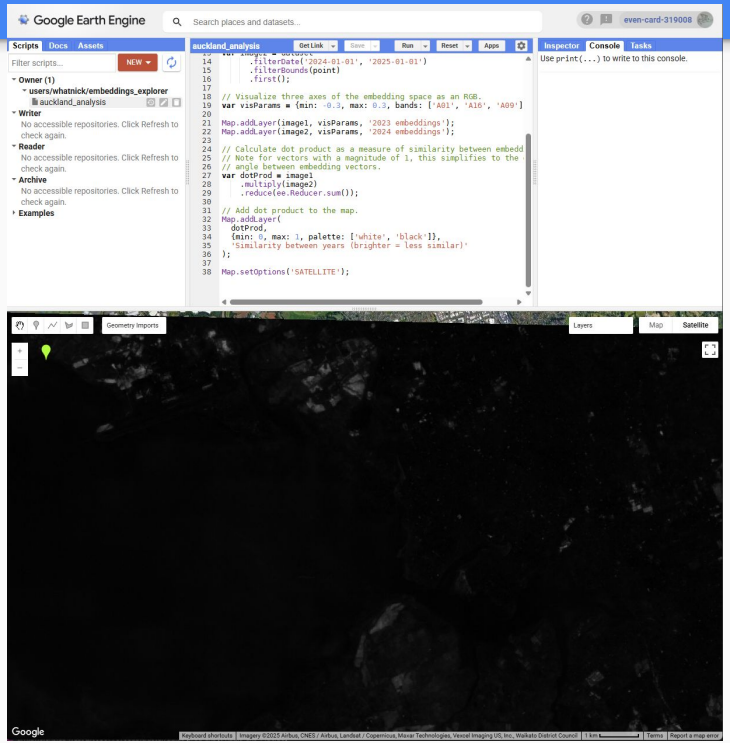
Image Credit : IBM - ViT + Location embeddings

Picking one and using it

- Technical Criteria
- Operational and Strategic Criteria
- Risk, Ethics, Governance and Data Provenance
- Task specific trade off - Accuracy vs Cost, Generality vs Specialization, Multimodality vs Simplicity, Proprietary vs Open-Source

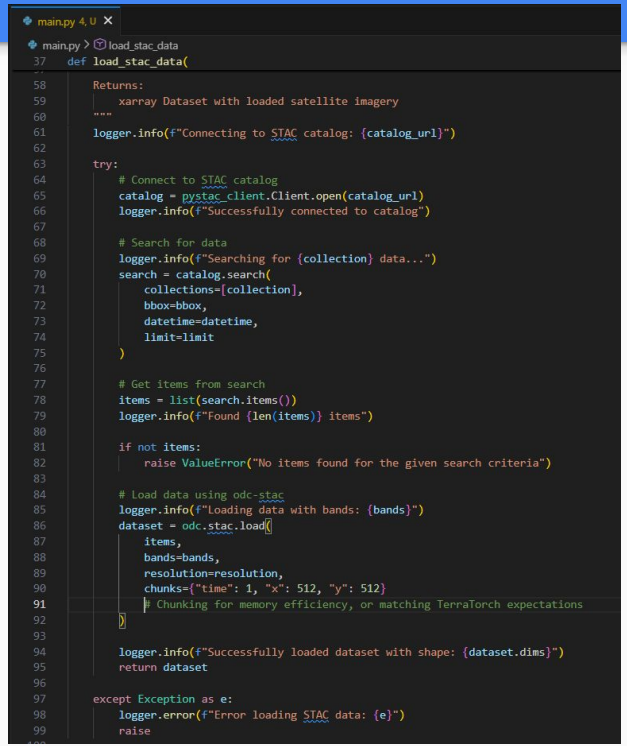
What about Embeddings ?

- Picking embeddings requires similar considerations like picking Foundation models
- Embeddings allow flexibility in mapping from per pixel representation to required usable modality



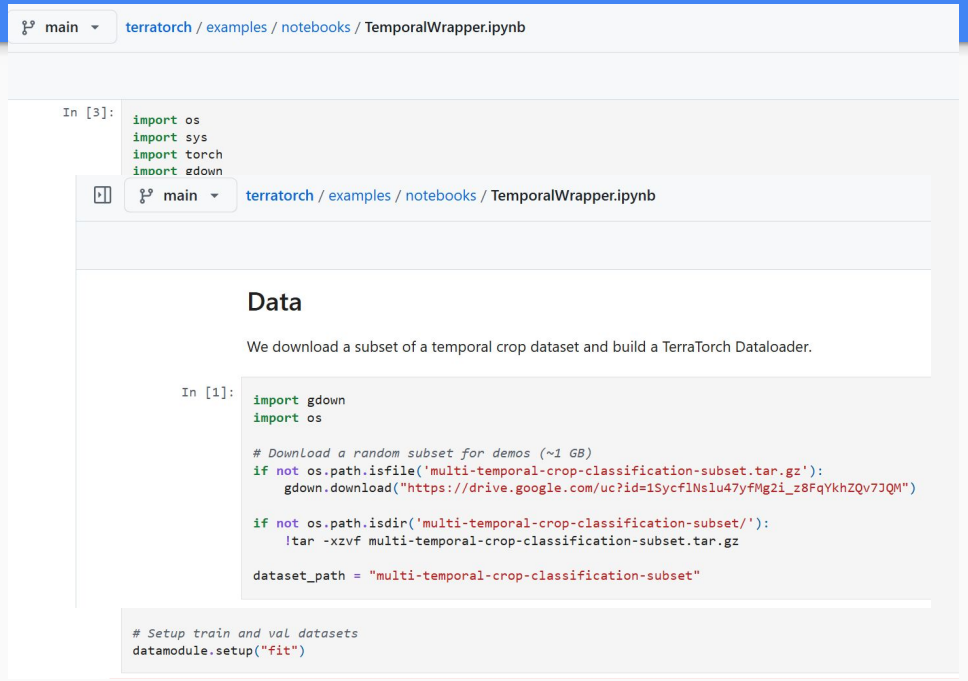
Loading some satellite imagery - STAC Way

- ODC-Stac is a convient way to load time series from COG (and soon Hyperspectral Zarrs).
- A bunch things behind STAC are actually HDF5 and NetCDF. More on that later.



Loading some satellite imagery - GeoFM Way

- GeoFM's typically implement Pytorch Derived DataLoaders
- These loaders focus on reformatting data for the ML models rather than focusing on cloud native access.



Picking a problem to solve - Modality

- Foundation models can be multi-modal i.e. they are useful simultaneously for solving multiple problems
- Picking a GeoFM requires consideration of which modalities it was originally trained for and whether the target modality is one of them or suitably adjacent to them e.g. a GeoFM covering Forestry modalities may be suitable for Grasslands too.
- Multi-modality with it often carries the possibility of confusion with different tasks if they are not well separated

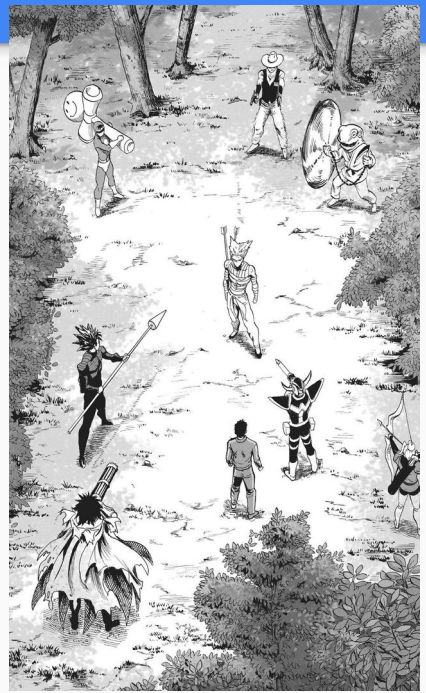


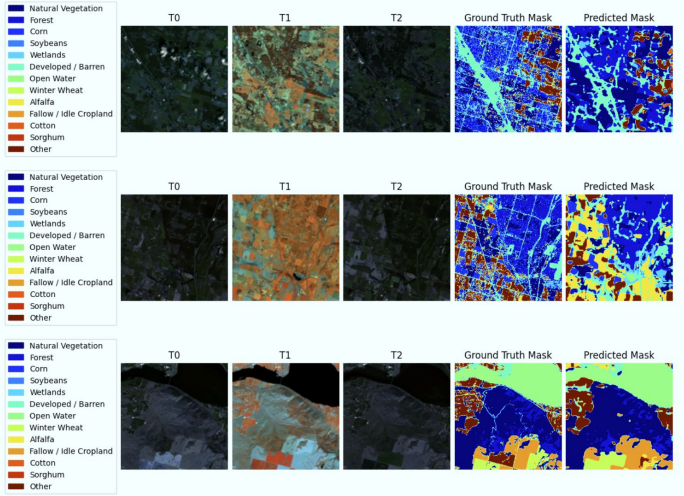
Image Credit : Yusuke Murata

Segmentation Modality

Both Prithvi-1.0 and 2.0 provide segmentation modalities which are useful for classification purposes at a pixel or object level.

In [19]:
```
# Get first batch of samples for plotting

data_loader = trainer.predict_dataloaders
batch = next(iter(data_loader))
```

In [20]:
```
# Plot samples and respective predictions

for i in range(BATCH_SIZE):

    sample = {key: batch[key][i] for key in batch}
    sample["prediction"] = preds[0][0][i].cpu().numpy()

    data_module.predict_dataset.plot(sample)
```



Regression Modality

- Regression involves inferring single or multiple parameters from source satellite imagery reflectance
- In quantitative data sparse cases regression modalities can be powerful in scaling parameter retrievals across space/time

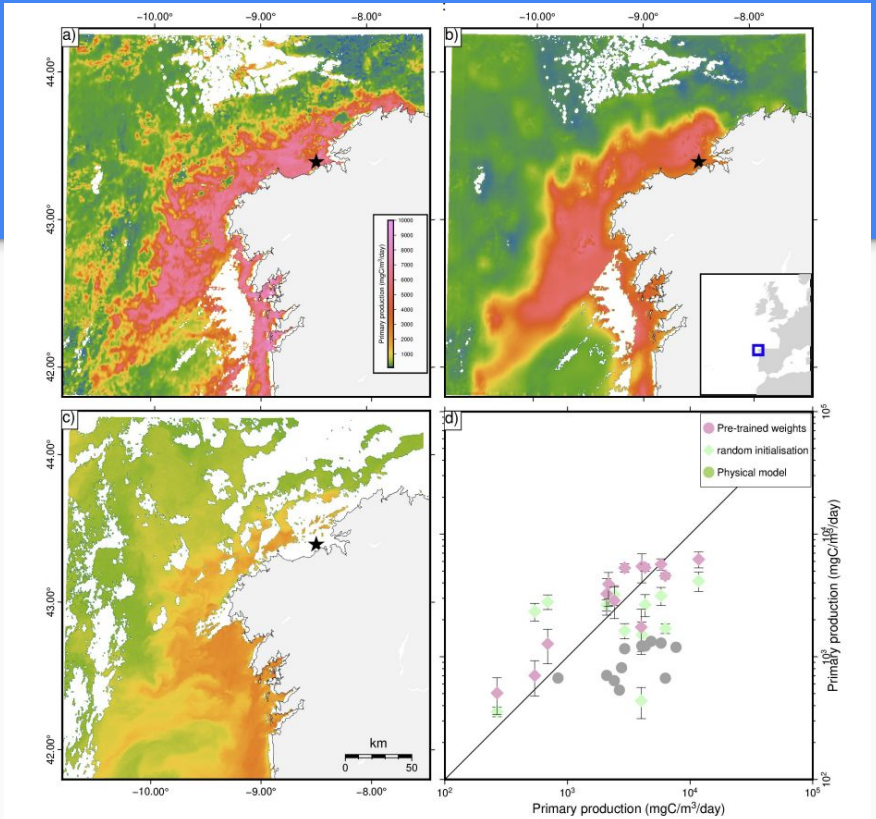


Figure 7: Primary production shown on log scale. (a) Inference from model fine-tuned on OCLI+SST; (b) inference with *scratch* model; (c) prediction of physical model; (d) comparison of models at measurement site (location indicated by a star on the spatial plots)

Source : [Granite Geospatial Ocean preprint]

Visual Question Answering (VQA) Modality

- Looking through satellite imagery especially after a disaster to assess damage is tedious and multiple questions may need answering
- A VQA style foundation model can rapidly answer multiple questions as they arise during the response process.

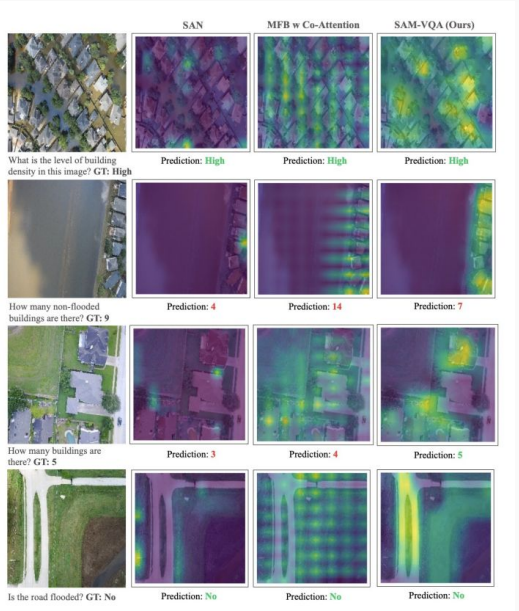


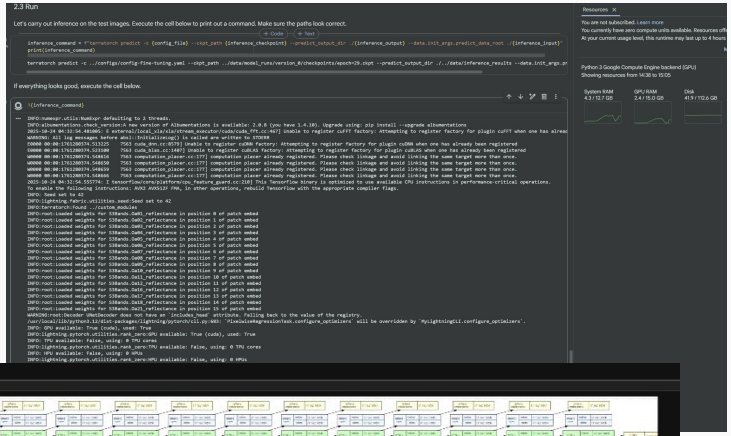
Image Credit : SAM-VQA: Supervised Attention-Based Visual Question Answering Model for Post-Disaster Damage Assessment on Remote Sensing Imagery

Some code Samples - Model Weights

- Loading a set of pretrained model weights and architecture
- Loading Data chip to perform inference on
- Loading Data chips to perform fine tuning on
- Iterations and loss-functions
- Inference and saving results

Some code Samples - Model Weights

- Loading a set of pretrained model weights and architecture
- Loading Data chip to perform inference on
- Loading Data chips to perform fine tuning on
- Iterations and loss-functions
- Inference and saving results



Distillation and Running on satellites

- Foundation models can be distilled with highest importance weights to be run at the edge.
- This opens up opportunities for in-space detection and regression for near-real time monitoring

Once is not enough - Continuous Training

- Embeddings have to be recomputed over collections regularly to assimilate new data
- Models need to be retrained regularly as new imaging options become available and as research in model architectures continues.
- Fine-tuning may have some undesired effect in shifting the subtle internal global optimum the model has achieved
- All this of-course requires a sustainable business model,.much like the rest of EO and robust down stream sources.

Opportunities in GeoFM Development - SAR

- Current State-of-the-Art (SOTA) models use amplitude only version of SAR data sets.
- Opportunities exist to incorporate cloud-native Geocoded Single-Look-Complex (GSLC) products being made available to incorporate phase information

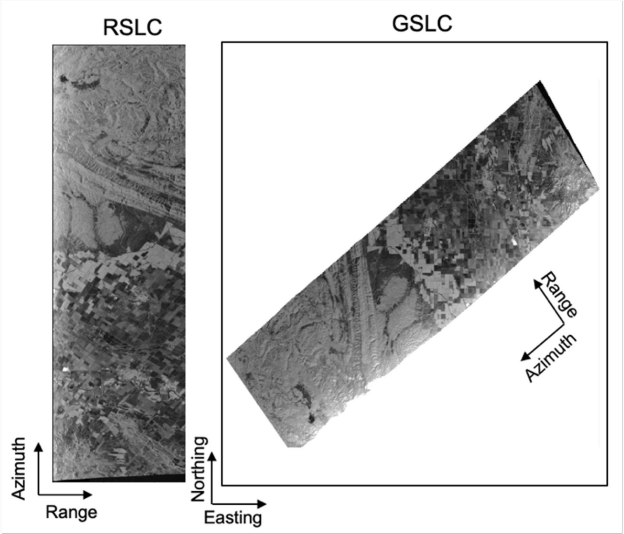
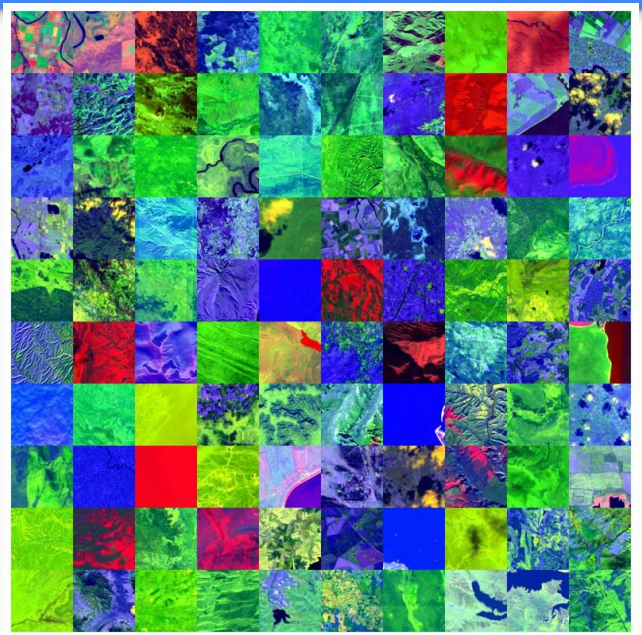


Image Credit : NISAR Product Sample from NASA/ISRO

Opportunities - HyperSpectral

- Hyperspectral Foundation models are not as prolific, but attempts and frameworks exist.
- The most common is EnMap based SpectralEarth. This is a combination of a training dataset and a foundation model.
- As more Hyperspectral data becomes available true foundation models will emerge.
- We need better data distribution process for HS for this to become a lower bar.

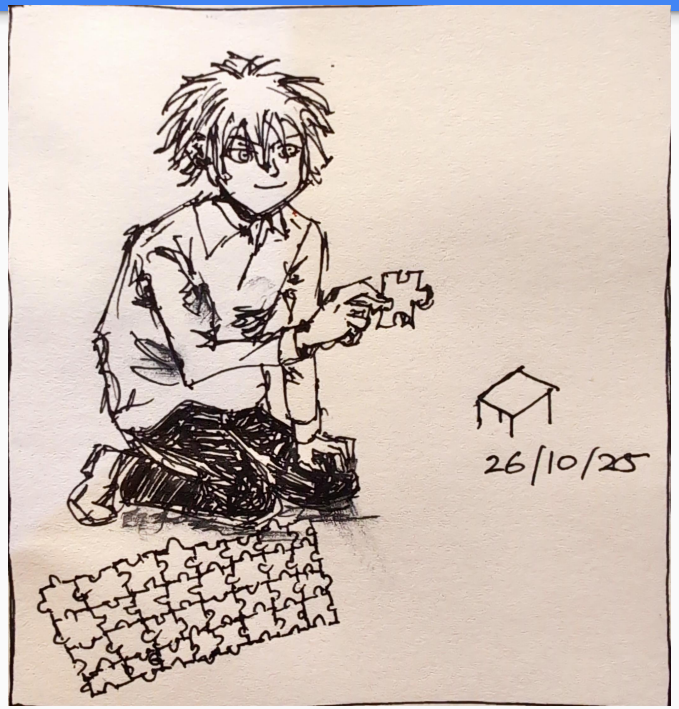


Opportunities - Power to Frozen Weights

- Crypto was seen as a way to dump excess power into something valuable. There are emerging trends for - Proof-of-Useful-Work which looks to store energy as frozen weights.
- In a process similar to creating ARD's , creating ARE's (Analysis Ready Embeddings) or DRM's (Deployment Ready Models) will be fashionable.
- Acronyms are a kind of frozen understanding

Opportunities - More ARCO and Dataloaders

- The world of Foundation models (and research into them) has to bridged by the world on Cloud-Native Geospatial and associated engineering.
- The FOSS4G community can looks at building solutions to wire leading GeoFM libraries to leading cloud native data access patterns - in languages of their choice.



Questions ?

https://github.com/whatnick/foss4g_2025_geofm

WHATNICK_INK

Tishampati Dhar

Separating complex and complicated