Radiant Earth Foundation

EARTH IMAGERY FOR IMPACT

Radiant MLHub: Advancing Utilization of AI Applications on Earth Observations with Benchmark Training Datasets

Hamed Alemohammad

Chief Data Scientist & Executive Director

Radiant Earth Foundation

2nd NOAA Workshop on "Leveraging AI in Environmental Sciences"

February 4, 2021

Mission

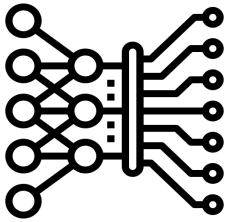
Empowering organizations and individuals globally with open Earth observation training data, standards and tools to cultivate a global community focused on machine learning and Earth observations to meet the world's most critical challenges.

Vision

Leveraging machine learning and Earth observation for positive global impact

Motivation

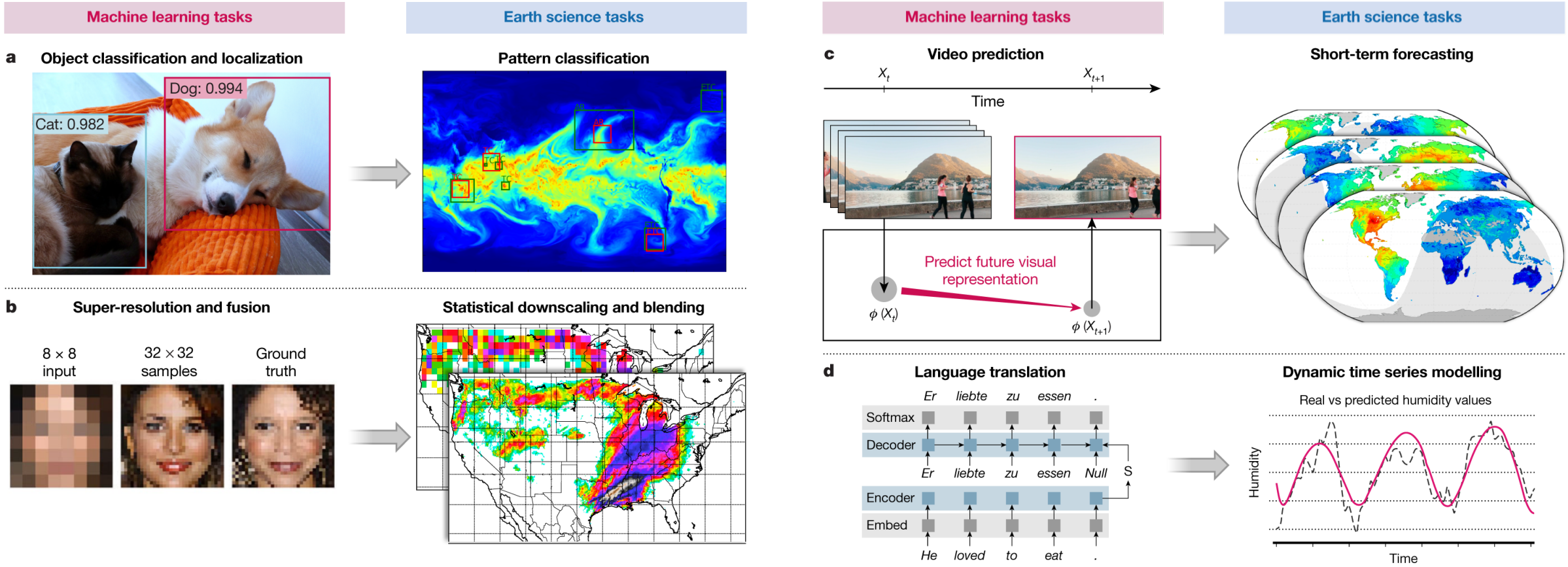
- **Increasing volume of data**
  - New satellites, in situ sensors, and models
  - Governments and commercial entities



- **Advancements in Machine Learning techniques**
  - Data-driven and fast iterations
  - Capable of detecting complex and non-linear relationships

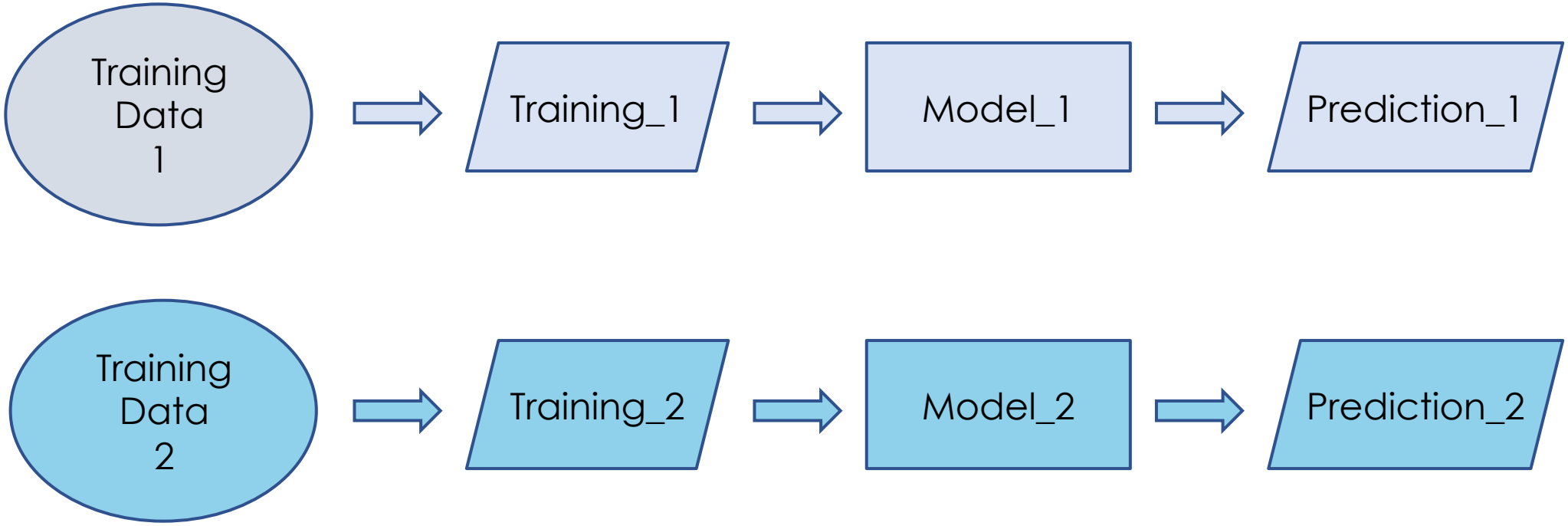
- **Availability of cloud services**
  - Brining computation to data
  - Scaling resources on-demand
  - Serverless designs

ML Tasks in Earth Science

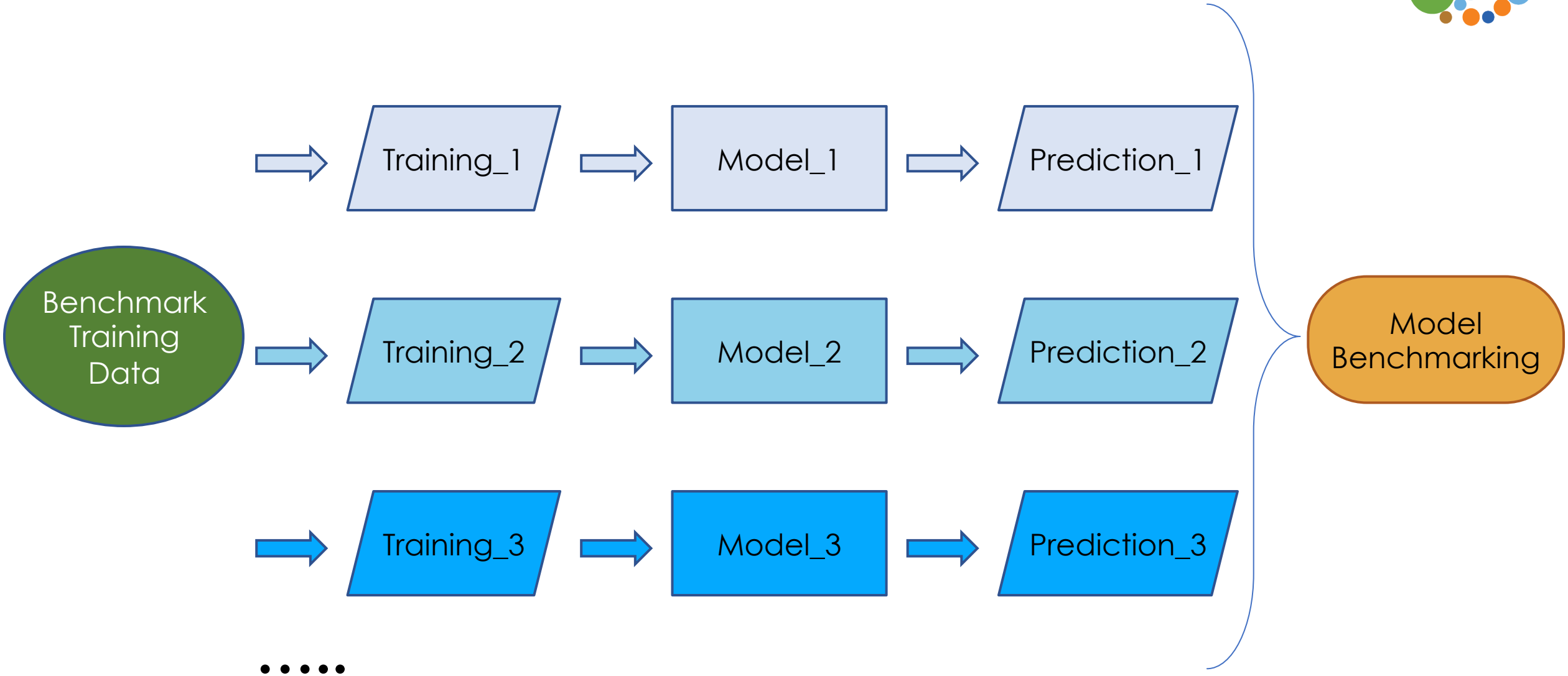


Source: Reichstein. et al., 2019

Existing Workflows

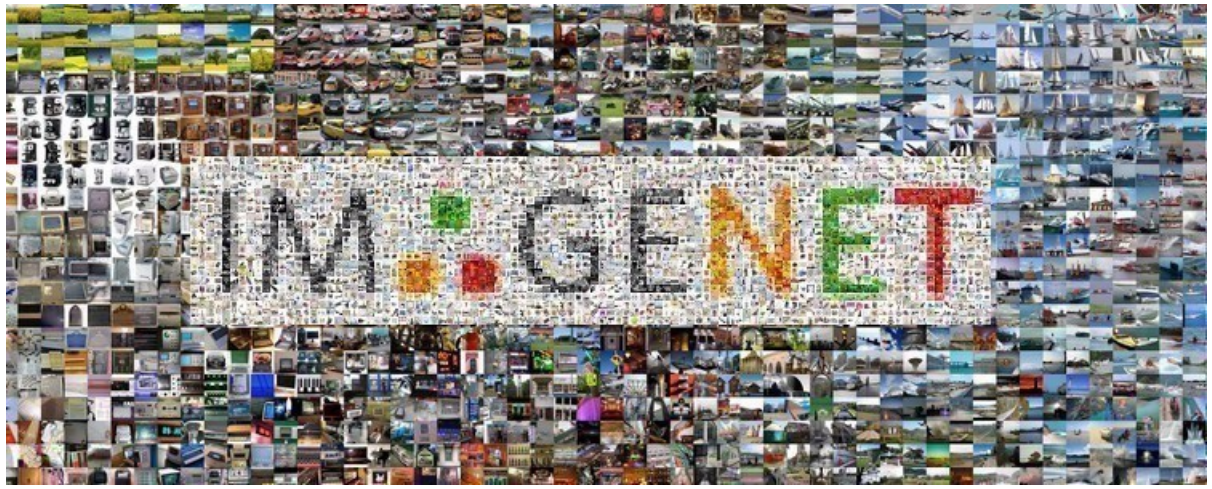


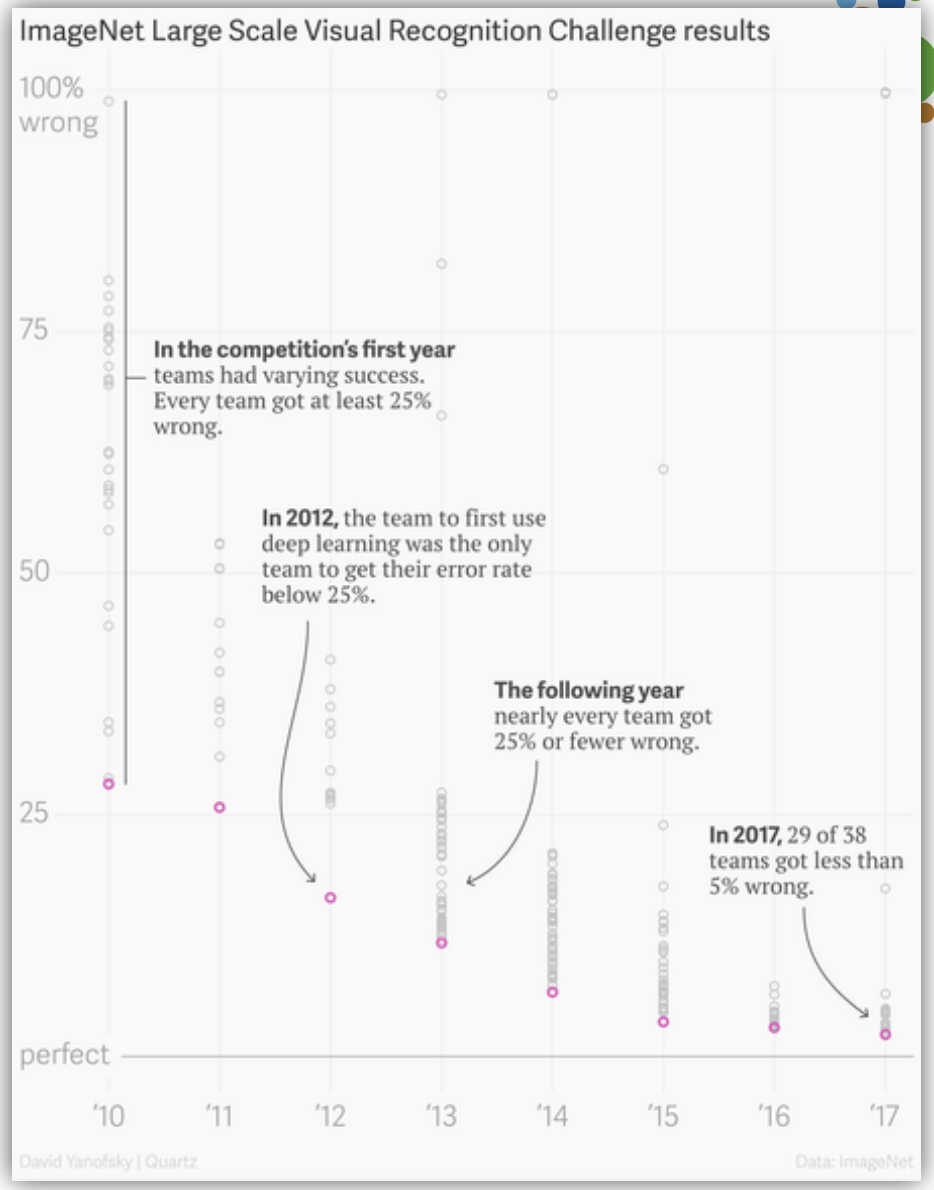
Ideal Workflow



ImageNet

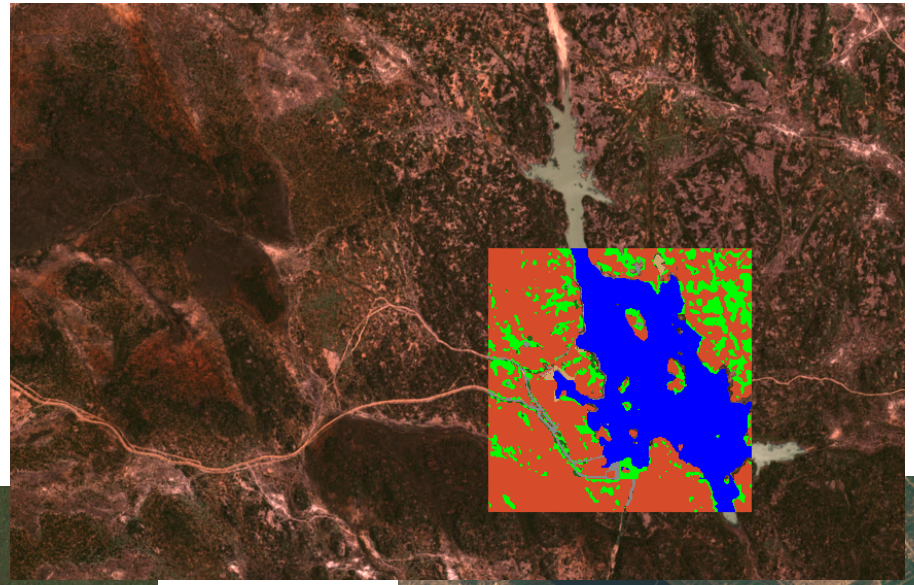
- 14 M annotated images including 1 M with object bounding boxes.
- 20 K categories of objects
- Open access
- Annual competition 2010-2017

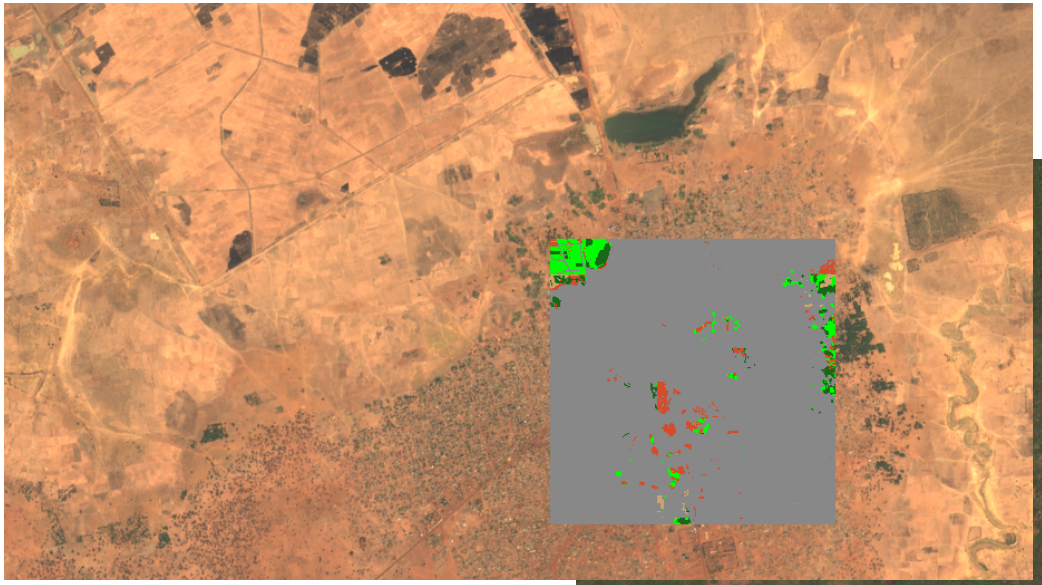


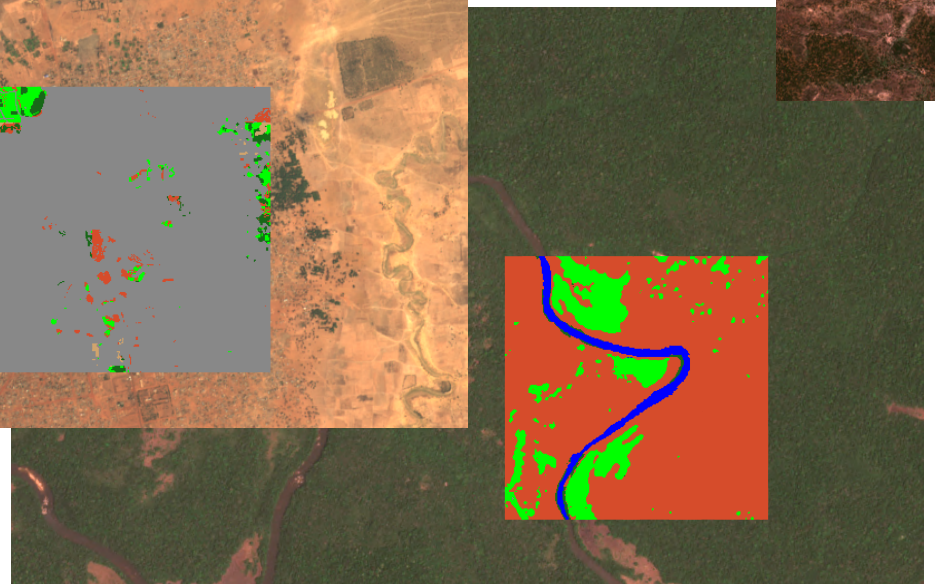


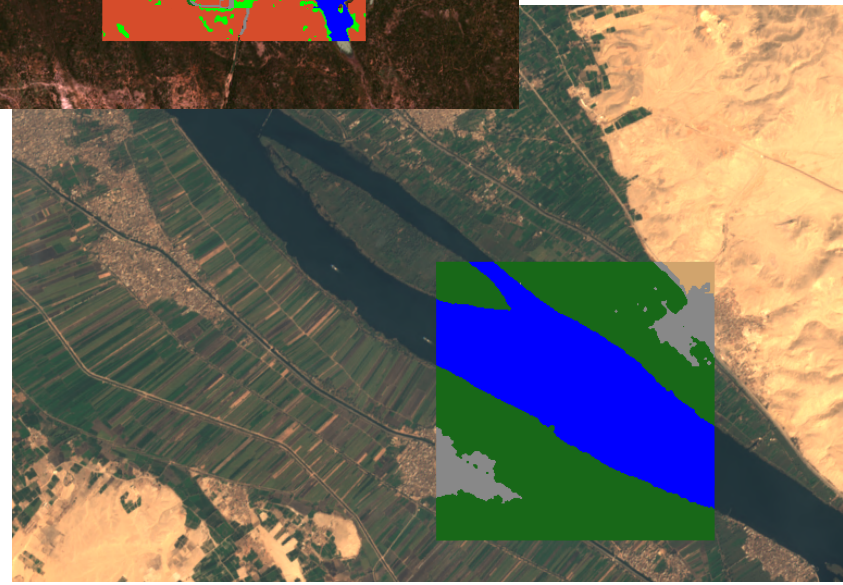
LandCoverNet

- First geodiverse LC training dataset
- 130M labeled pixels at 10 m resolution
- 7 classes based on annual time series









landcover.net

WeatherBench

- A benchmark dataset for data-driven weather forecasting (based on ERA-5)
- 14 variables including
  - Temperature
  - Relative Humidity
  - Vorticity
  - Precipitation
  - Cloud Cover
  - ….
- Evaluation metrics and baseline models

GitHub repo

Training Data Challenges in Earth Science

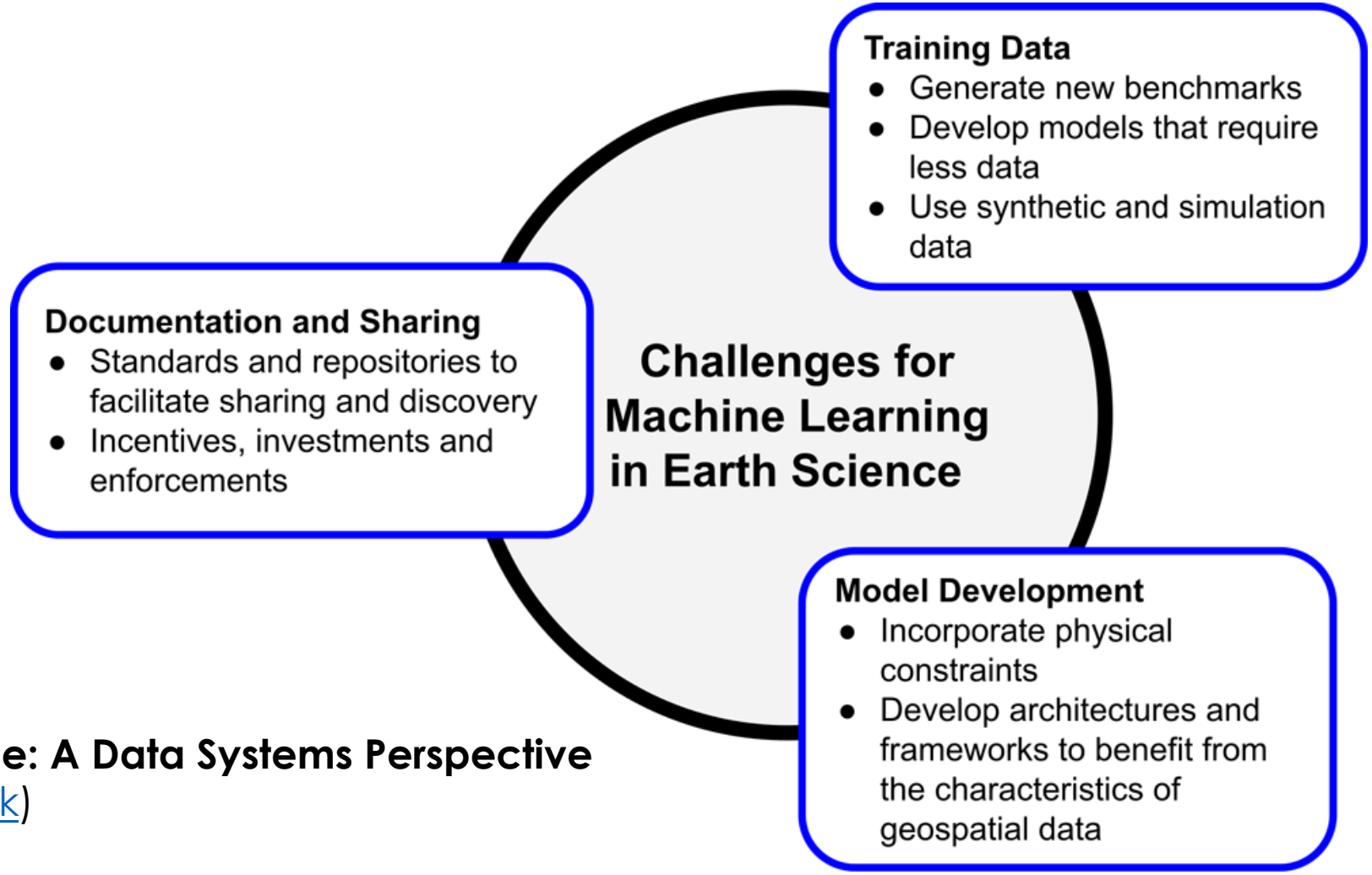
Geospatial Training Data Catalogs:

- Lack of Geo-Diversity
- Scarce data sources
- Data Accessibility
- Inter-Operability
- Machine learning-readiness

Result of Gaps in Training Data Catalogs:

- Biased or incorrect results
- Inability to capture wide range of possible outcomes in space and time

Challenges for ML in Earth Science



Advancing AI for Earth Science: A Data Systems Perspective
AGU EOS, November 2020 (link)

ML Commons for Earth Observation

Hub

- EO Training Datasets
- ML Models
- Competitions
- Image annotation + ground-referencing

Community

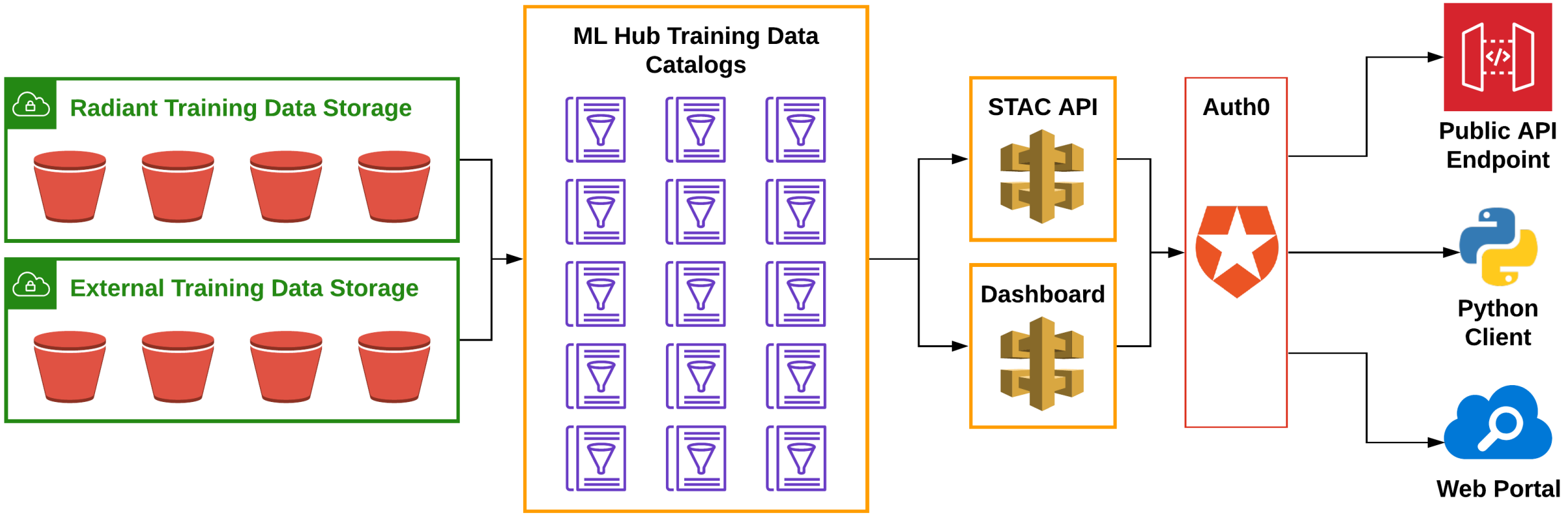
- Convenings to develop standards for ML on EO
- Interoperability of datasets
- Technical Working Groups
- White Papers

Education

- EO market information
- Best practices on use of ML and EO
- Speaking engagements
- Media outreach

Radiant MLHub
Earth Imagery for Impact

Radiant ML Hub



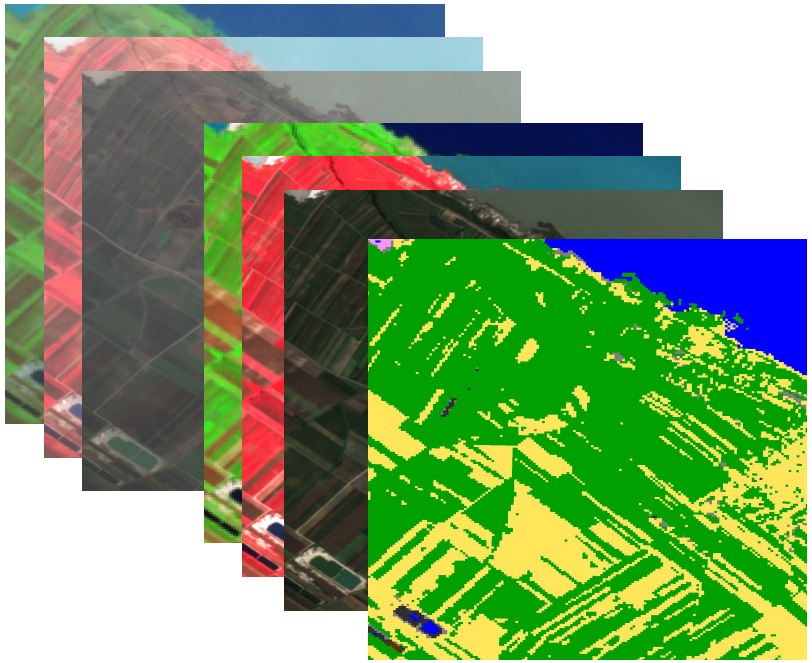
Radiant MLHub Repository

- Each dataset has a DOI with version and citation

- FAIR data principles
  - Findable
  - Accessible
  - Interoperable
  - Reusable

Radiant MLHub Data Catalog

- **Crops**
  - Uganda
  - Kenya
  - Tanzania
  - South Sudan
  - Ghana
  - Mali
  - South Africa
  - Rwanda
  - Central Asia
  - Benin
- **Land Cover**
  - LandCoverNet
  - BigEarthNet (EU)
  - SpaceNet 1-6
  - Chesapeake Bay
  - SpaceNet 7
- **Building Footprints**
  - Tanzania
- **Tropical Storms**
  - Pacific and Atlantic
- **Wildfire Fuel**
  - Western US
- **Surface Water**
  - Multiple regions



Competitions

Past:

- Crop type classification in Kenya

  Multiband and temporal Sentinel-2

- Cloud labeling

  Optical data from Sentinel-2

- Tropical Cyclone Wind Speed Estimation

  Temporal observations from GOES

Future:

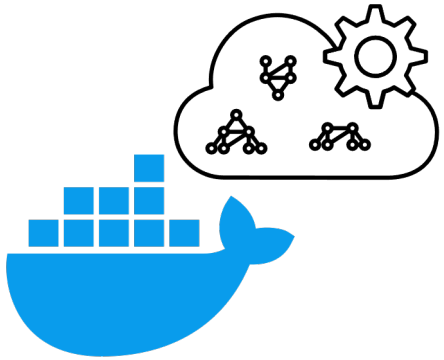
- Crop Monitoring in South Africa (Summer)

  Multiband and temporal Sentinel-2 and Sentinel-1

What's coming up?

- Enhance training data repository
  - Develop a Python Client (alpha version released)
  - Build integration with NASA EOSDIS CMR catalog

- Develop an ML model repository
  - Commit your model in GitHub
  - Register on Radiant MLHub
  - Share with the world!



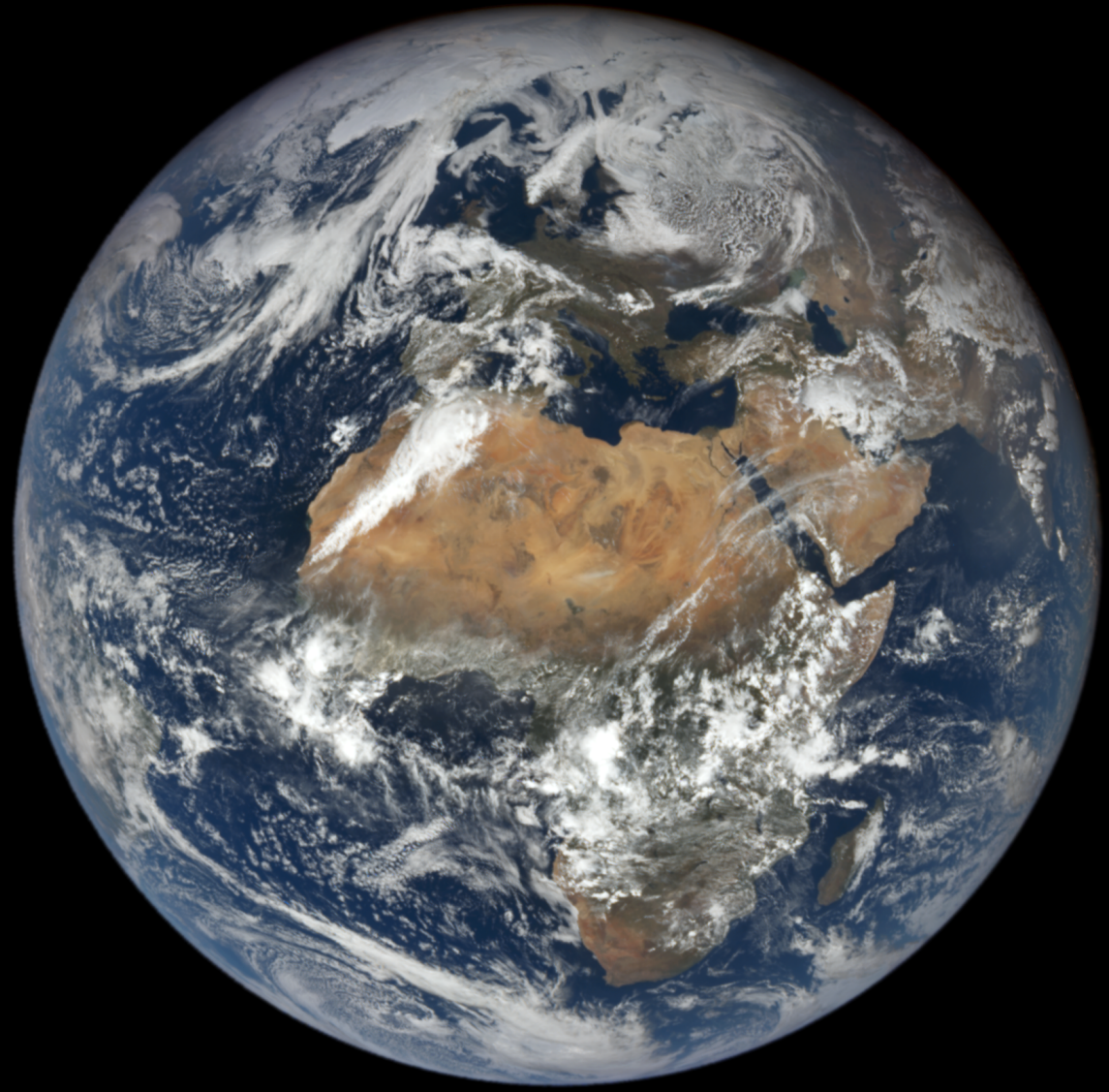
- Generate new training datasets
  - Fusion of SAR and multi-spectral data

Thanks!

www.radiant.earth

www.mlhub.earth

MLHub Slack Channel:
bit.ly/MLHubSlack



Funding Support

OMIDYAR NETWORK

BILL & MELINDA GATES foundation

NASA Partner

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

SCHMIDT FUTURES

Patrick J. McGovern FOUNDATION