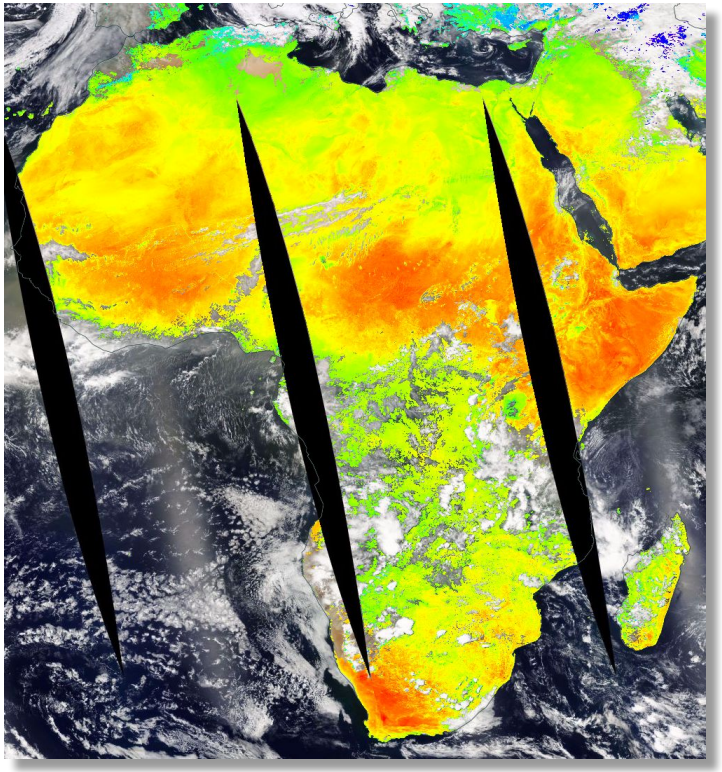
NASA briefing to Unidata

Dr. Christopher Lynnes[1] and Dr. Christine Smit[2]

[1]NASA Earth Science Data and Information Services (Civil Servant)
[2]NASA Goddard Earth Sciences Data and Information Services Center (Telophase)

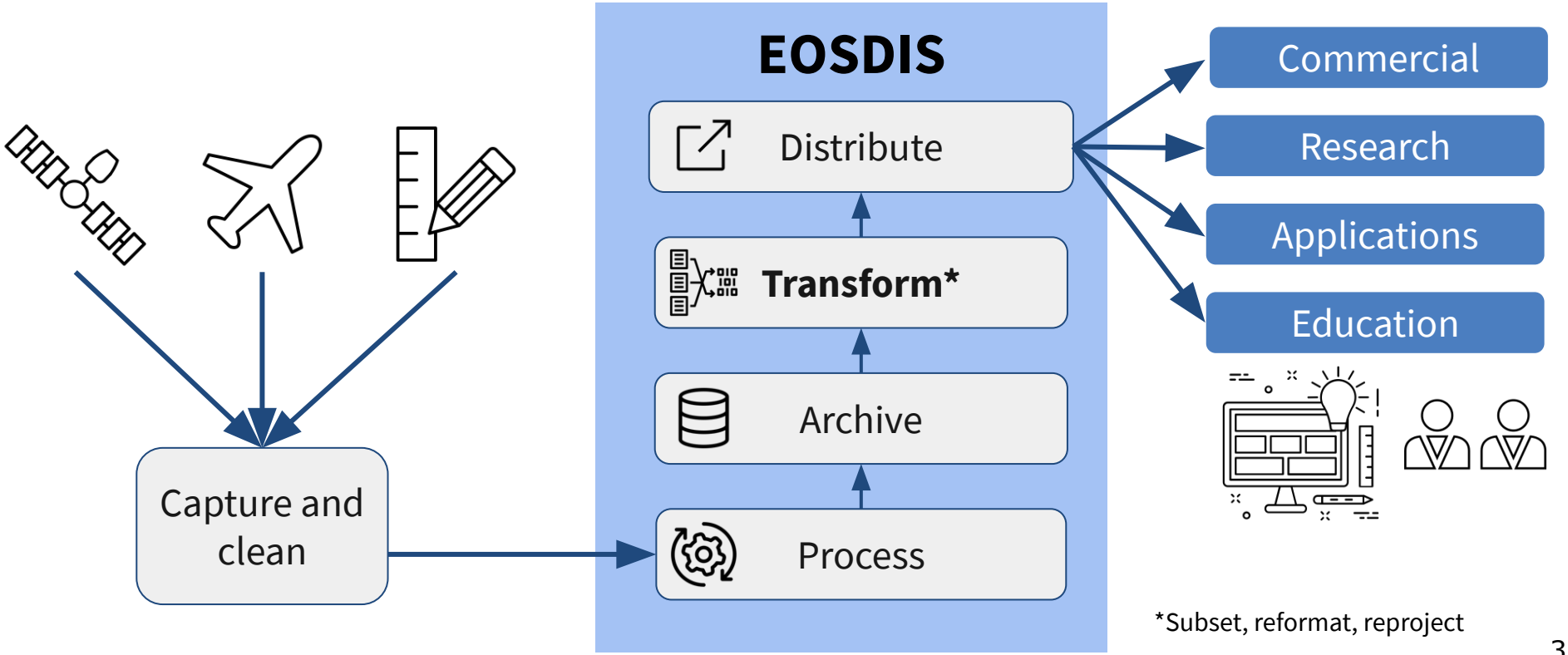
NASA's Earth Science Data Systems Program

- Actively manages NASA's Earth Observations
  - Multivariate
  - Heterogeneous
  - Diverse satellites and suborbital platforms
  - Most spheres: atmo, bio, cryo, hydro, litho
- Designed to support rigorous science **research**
- Processes instrument data to create high quality *long-term* Earth science data records.



http://go.nasa.gov/2mMd5g1

Earth Observing System Data and Information System (EOSDIS)



*Subset, reformat, reproject

Why Are We Here?

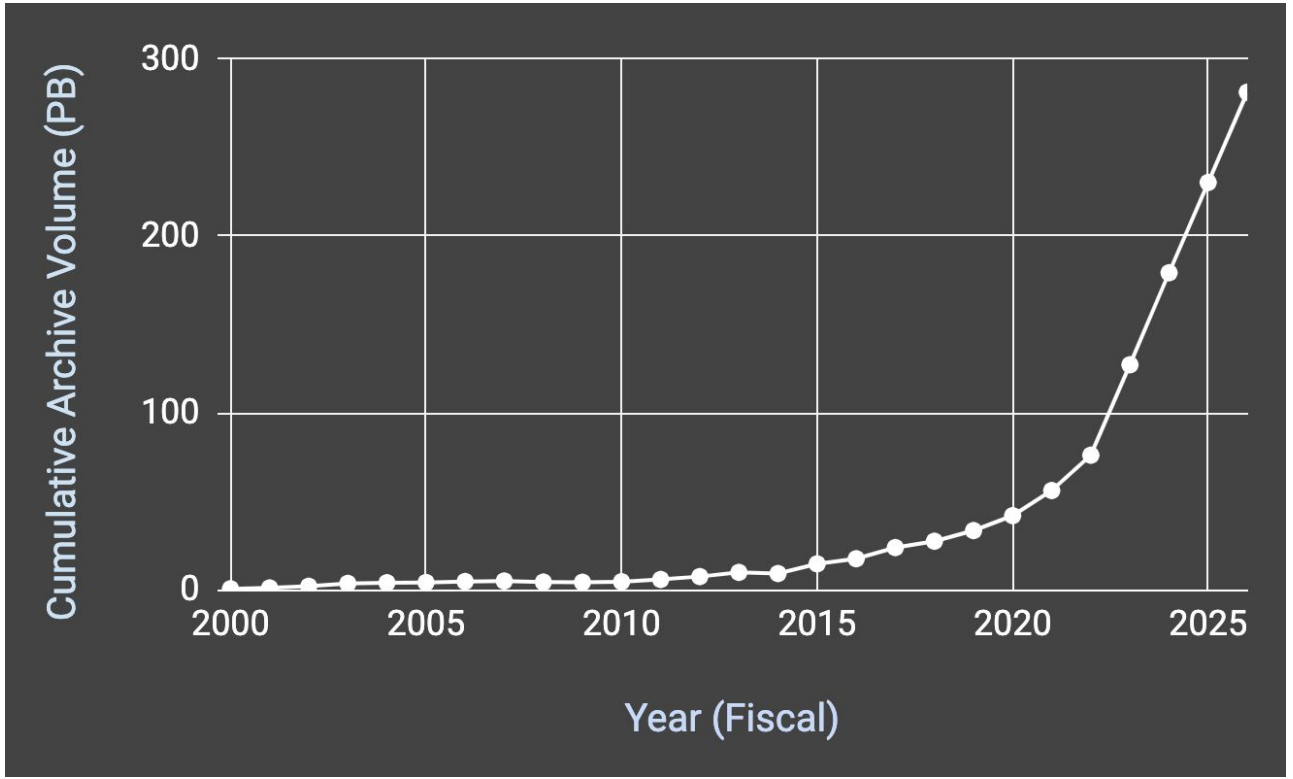
1. NetCDF[1]
   a. Recommended EOSDIS format for gridded data
   b. At least 1798 datasets
2. THREDDS[2] Data Server
   a. Oak Ridge National Laboratory DAAC[3]
   b. Physical Oceanography DAAC[3]
   c. Goddard Earth Sciences Data and Information Services Center

[1]network Common Data Form
[2]Thematic Real-time Environmental Distributed Data Services
[3]Distributed Active Archive Center

EOSDIS in the Cloud

Downloading data to local computers will become unwieldy as data volumes explode.

EOSDIS is migrating data to commercial cloud to allow users to utilize Big Compute next to Big Data.





Thank you!

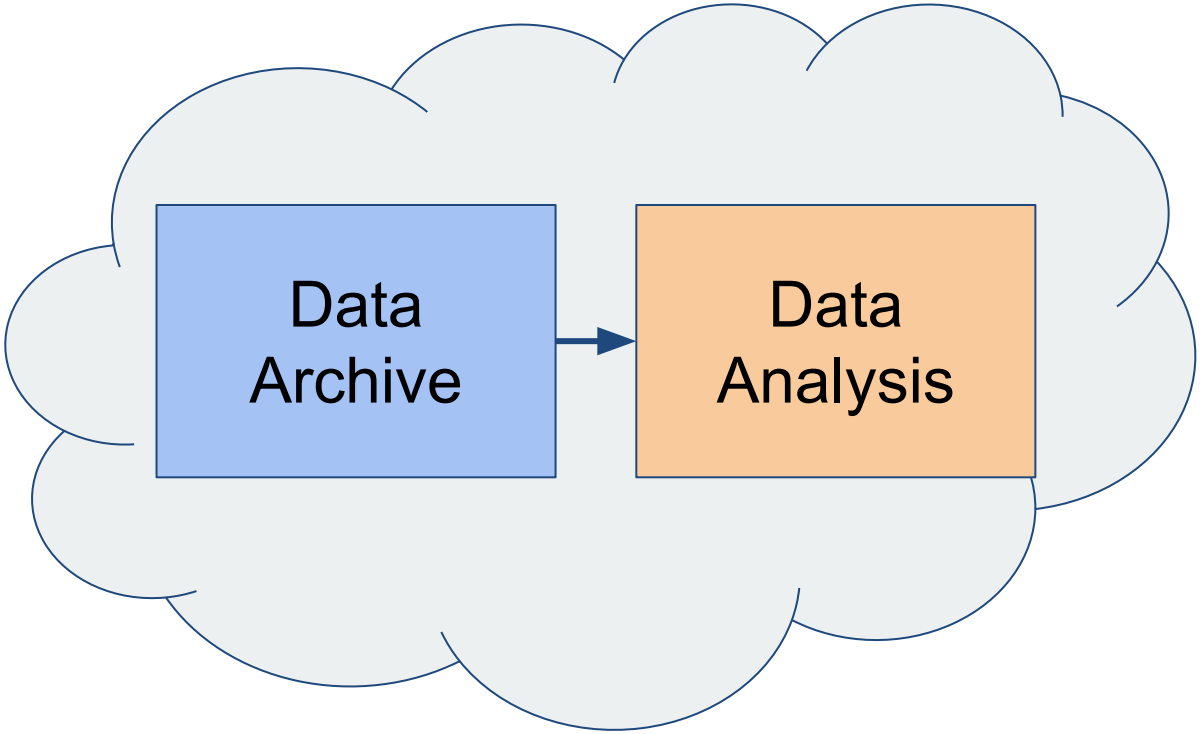
Thanks to

Ward, Sean, Dennis, and Ethan

for the excellent webinar in April on

NetCDF in the Cloud

Analysis-in-Place



Data Migration to AWS[1] Cloud

- 866 datasets in the cloud now:
  - Global Hydrology Resource Center: Severe Weather
  - Alaska Satellite Facility: Synthetic Aperture Radar
  - Physical Oceanography DAAC: Sea Surface Temperature et al.
  - Land Processes DAAC:
    - Harmonized Landsat/Sentinel-2
    - ASTER[2] Digital Elevation Model
- Planning to accelerate migration...

[1] Amazon Web Services
[2] Advanced Spaceborne Thermal Emission and Reflection Radiometer

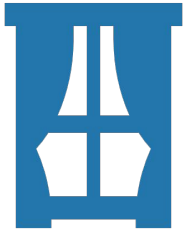
Enabling Analysis-in-Place

- Transformation services to make analysis-ready data
  - Some conversion to cloud friendly formats (zarr, CoG[1])
  - Spatiotemporal Asset Catalog (STAC) delivery manifests
  - OPeNDAP[2] Hyrax and THREDDS Data Server (in work)
- In Work: Small but Critical User Experience Aspects
  - Token-based authentication
  - Latency
  - Exposing S3[3] links
  - Instructions on in-place usage

[1] Cloud-optimized Geographic Tagged Image File Format
[2] Open-source Project for a Network Data Access Protocol
[3] Simple Storage Service

The Big News: Open Source Science

The Four Meanings of “Open” in Open-source Science



OPEN (**TRANSPARENT**) SCIENCE
scientific process and results should be visible, accessible, and understandable

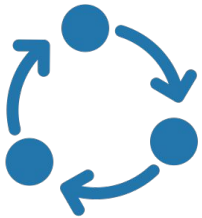
OPEN **(ACCESSIBLE)** SCIENCE
data, tools, software, documentation, and publications should be accessible to all





OPEN **(INCLUSIVE)** SCIENCE
process and participants should welcome participation by and collaboration with diverse people and organizations

OPEN (**REPRODUCIBLE**) SCIENCE
scientific process and results should be open such that they are reproducible by members of the community



Building an Open Sourced Science Ecosystem

- Initiate new missions, research and applied activities as open science projects.
- Implement clear policies for software, publications and data.
- Integrate and improve existing capabilities to support data management, access, computing, analytics and collaboration.
- Build the community through training, workshops, competitions and incentives.



Transform to OPen Science (TOPS) is a 5-year effort focused on capacity building, partner engagement, and incentives to help accelerate scientific discovery through open science.

TOPS Focus Areas

Public Engagement	• **Designate 2023 as Year of Open Science** • Partnering with professional orgs., publishing TOPS articles in high-impact journals • Engage early with historically excluded communities • TOPS GitHub - https://github.com/nasa/Transform-to-Open-Science
Capacity Building	• Create FAIR* - Analysis-Ready Cloud-Optimized (ARCO) data • Develop learning resources • TOPS JupyterHub • Host and sponsor events (summer schools, multi-day trainings, massive open online courses)
Incentives	• Develop NASA Open Source Science Awards program • Leverage prizes and challenges and cross-division science use cases • Increased citizen science activities

*Findability, Accessibility, Interoperability, and Reusability