



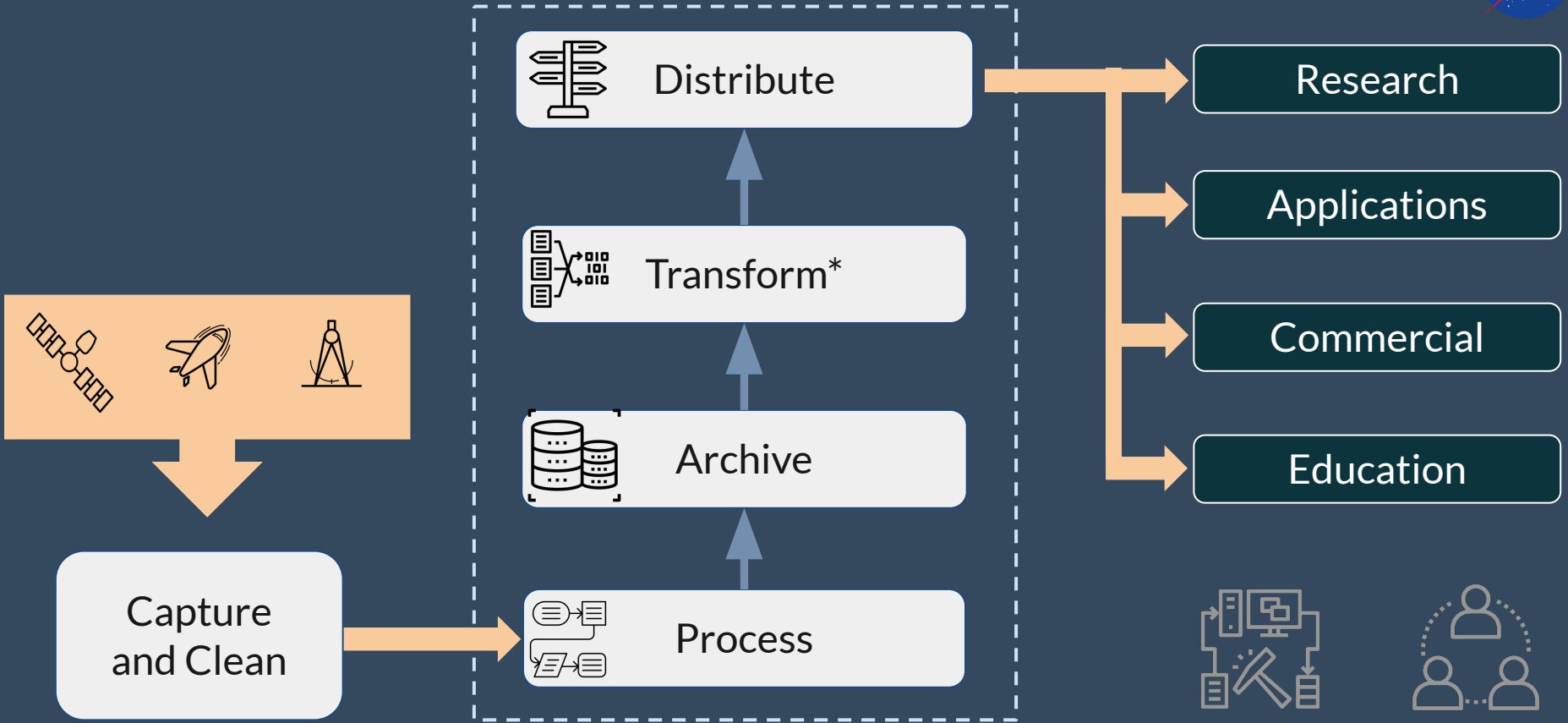
Federated Cloud Challenges in NASA's Earth Science Data Systems (Why So Difficult?)

Christopher Lynnes¹ and Katie Baynes¹
EOSDIS² System Architects

¹ NASA Civil Servant

² Earth Observing System Data and Information System

Earth Observing System Data and Information System



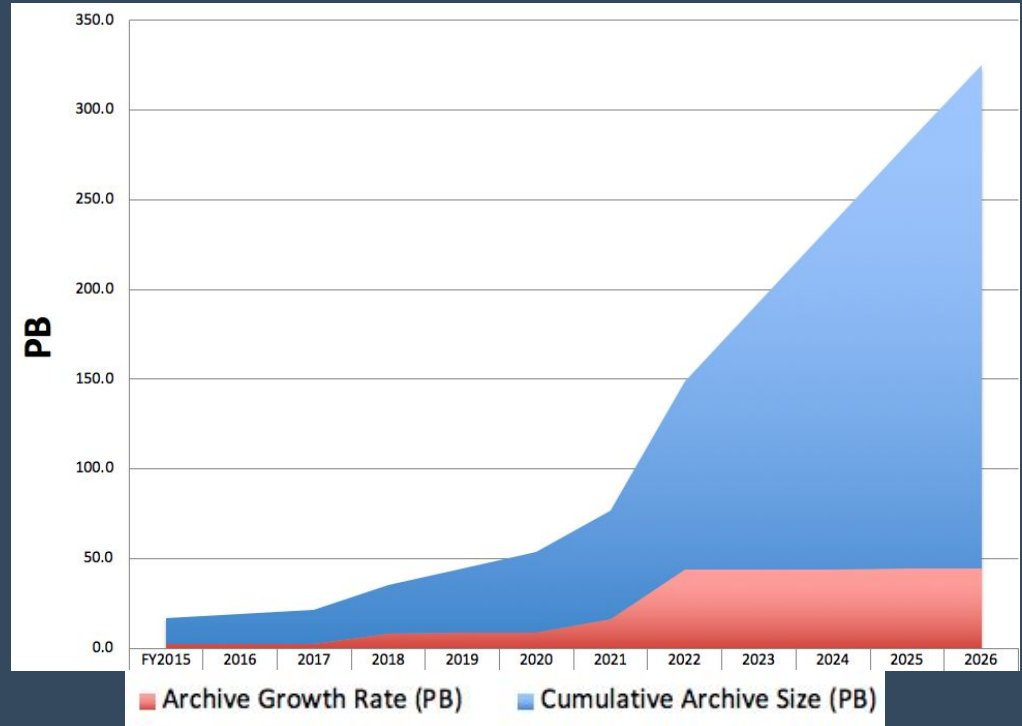
*Subset, reformat, reproject

Using Cloud to Handle Big Data



The current on-premise architecture will be problematic as annual ingest rate increases from 4 to 45 PB/year as new missions begin operations.

ESDIS is developing open source cloud-native systems for reuse across the agency (and elsewhere).



Cloud offers the ability to *analyze data at scale, analyze multiple data sets together easily and avoid moving large data sets* allowing scientists, instead working on data “in place”

The “Other” V: Variety



Earthdata Search

NASA (National Aeronautics and Space Administration) [US] | https://search.earthdata.nasa.gov/search

EARTHDATA Find a DAAC

EARTHDATA Search

Browse Collections

Features

- Map Imagery
- Near Real Time
- Customizable

Keywords

Platforms

Instruments

Organizations

Projects

Processing levels

Type any topic, collection, or place name

6,099 Matching Collections

Sort by: Relevance Only include collections with granules Include non-EOSDIS collections

Tip: Add + collections to your project to compare and download their data.

Global Maps of Atmospheric Nitrogen Deposition, 1860, 1993, and 2050

27 Granules • 1860-01-01 to 2050-12-31 • This data set provides global gridded estimates of atmospheric deposition of total inorganic nitrogen (N), NHx (NH3 and NH4+), and NOy (all oxidized forms of nitrogen other than N2O), in mg N/m2/year, for the years 1860 and projections for the year 2050. The data set was...

1860_1993_2050_NITROGEN_830_v1 - ORNL_DAAC

IRS 1C LIS3 Standard Products

Int'1 / Interagency • 1996-11-14 to 2007-09-20 • The medium resolution multi-spectral sensor, LISS-3 operates in four spectral bands - B2, B3, B4 in visible near infrared (VNIR) and B5 in Short Wave Infrared (SWIR) providing data with 23.5m resolution. Standard products are full scene (path-row) based geo-referenced as well as...

v1.93.2 • Search Time: 0.6s • NASA Official: Stephen Berrick • FOIA • NASA Privacy Policy • USA.gov

Earthdata Access: A Section 508 accessible alternative

The “Other” V: Variety



Earthdata Search

NASA (National Aeronautics and Space Administration) [US] | <https://search.earthdata.nasa.gov/search>

EARTHDATA Find a DAAC

EARTHDATA Search

Type any topic, collection, or place name

6,099 Matching Collections

Browse Collections

- Features
 - Map Imagery
 - Near Real Time
 - Customizable
- Keywords
- Platforms
- Instruments
- Organizations
- Projects
- Processing levels

Sort by: **Relevance** Only include collections with granules Include non-EOSDIS collections

Tip: Add + collections to your project to compare and download their data.

Global Maps of Atmospheric Nitrogen Deposition, 1860, 1993, and 2050

27 Granules • 1860-01-01 to 2050-12-31 • This data set provides global gridded estimates of atmospheric deposition of total inorganic nitrogen (N), NHx (NH3 and NH4+), and NOy (all oxidized forms of nitrogen other than N2O), in mg N/m2/year, for the years 1860 and 1993 and projections for the year 2050. The data set was...

1860_1993_2050_NITROGEN_830_v1 - ORNL_DAAC

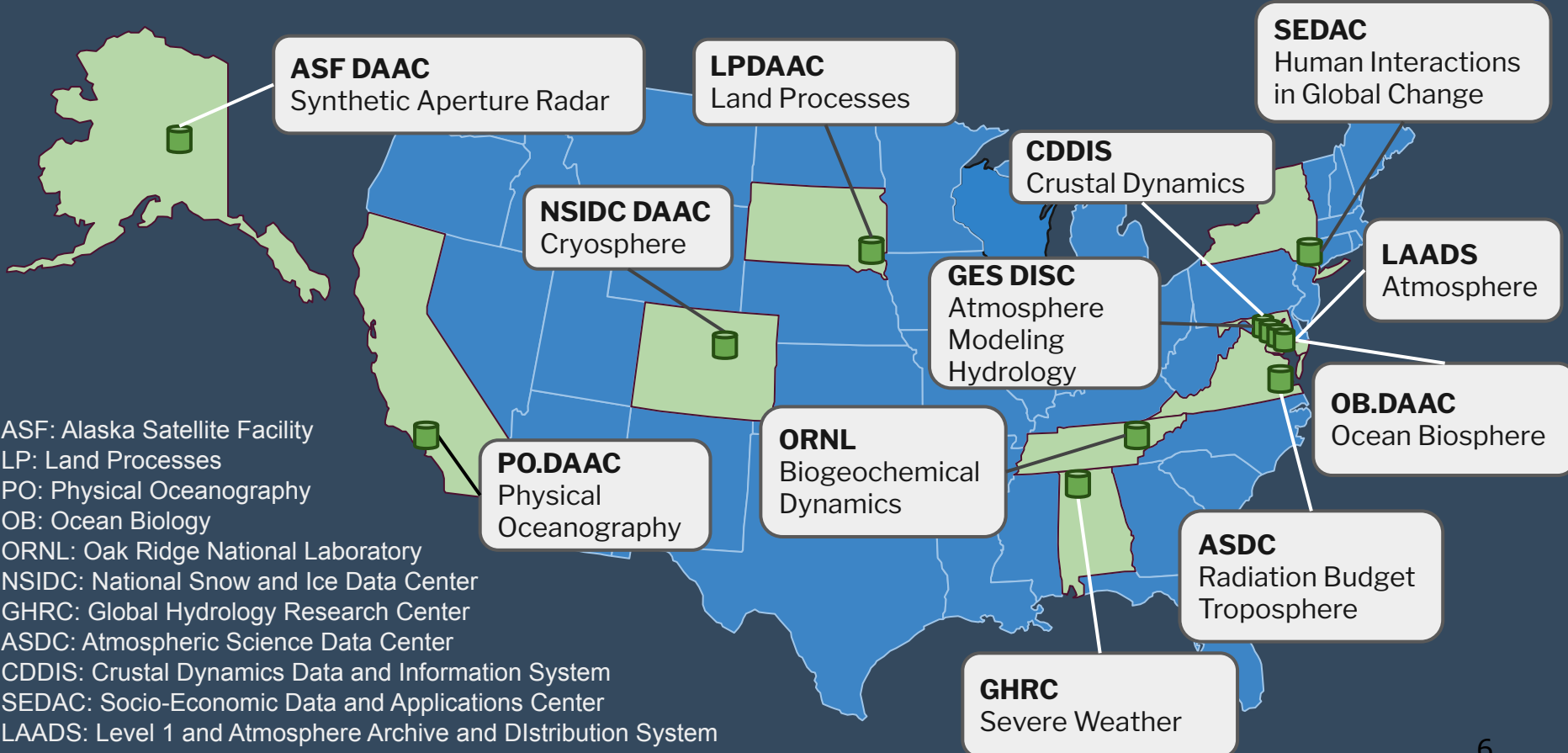
IRS 1C LIS3 Standard Products

Int'l / Interagency • 1996-11-14 to 2007-09-20 • The medium resolution multi-spectral sensor, LISS-3 operates in four spectral bands - B2, B3, B4 in visible near infrared (VNIR) and B5 in Short Wave Infrared (SWIR) providing data with 23.5m resolution. Standard products are full scene (path-row) based geo-referenced as well as...

v1.93.2 • Search Time: 0.6s • NASA Official: Stephen Berrick • FOIA • NASA Privacy Policy • USA.gov

Earthdata Access: A Section 508 accessible alternative

Distributed Active Archive Centers (DAACs)



ASF: Alaska Satellite Facility
LP: Land Processes
PO: Physical Oceanography
OB: Ocean Biology
ORNL: Oak Ridge National Laboratory
NSIDC: National Snow and Ice Data Center
GHRC: Global Hydrology Research Center
ASDC: Atmospheric Science Data Center
CDDIS: Crustal Dynamics Data and Information System
SEDAC: Socio-Economic Data and Applications Center
LAADS: Level 1 and Atmosphere Archive and Distribution System
GES DISC: Goddard Earth Sciences Data and Informations Services Center

DAACs and -Spheres of Influence



DAAC	Atmo-	Hydro-	Bio-	Cryo-	Litho-	Anthropo-
Atmospheric Sciences Data Center	✓					
L1 and Atmosphere Archive & Distribution System	✓					
Goddard Earth Sciences DISC	✓	✓				
Global Hydrology Resource Center		✓				
Physical Oceanography DAAC		✓				
Ocean Biology DAAC		✓	✓			
Oak Ridge National Laboratory DAAC			✓			
Land Processes DAAC			✓			
National Snow & Ice Data Center DAAC				✓		
Alaska Satellite Facility				✓	✓	
Crustal Dynamics Data Information System					✓	
Socioeconomic Data Archive Center						✓

How EOSDIS Federates



Search	Common Metadata Repository Unified Metadata Model
Browse	Global Imagery Browse System
Data	Standard formats and conventions Cloud storage
Services	TBD* (Web Coverage Service?)
Metrics	EOSDIS Metrics System
Sign-on	Earthdata Login (oauth2)
Budget	Cloud Tamer Egress cutoff and shaping

*TBD: To Be Determined



NASA - European Space Agency: MAAP

Multi-Mission Algorithm and Analysis Platform (MAAP)

- *Joint NASA-ESA Project*
- Biomass: Synthetic Aperture Radar + LIDAR*
- Key Features
 - Data analysis
 - Algorithm development
 - Science data processing
 - Collaborative capabilities

*Light Detection and Ranging,



How MAAP Federates

Search	Common Metadata Repository OpenSearch
Browse	Web Map Tiled Service
Data	Standard formats and conventions
Services	TBD
Analysis	TBD
Sign-on	OAuth2
Data Processing	TBD
Science Results	TBD

Committee of Earth Observing Satellites (CEOS)



- 62 Agencies
- 121 Satellites
- 5 Working Groups
 - Including Working Group on Information Systems and Services (WGISS)



How CEOS Federates

Search	OpenSearch
Analysis-Ready Data	Open Data Cube
Metrics	TBD

Group on Earth Observations (GEO)



- 105 Members
- 129 Participating Organizations
- GEOSS - Global Observation System of Systems

Future Federation in GEOSS



1. Systems
2. Knowledge
3. Data
4. Services
5. People

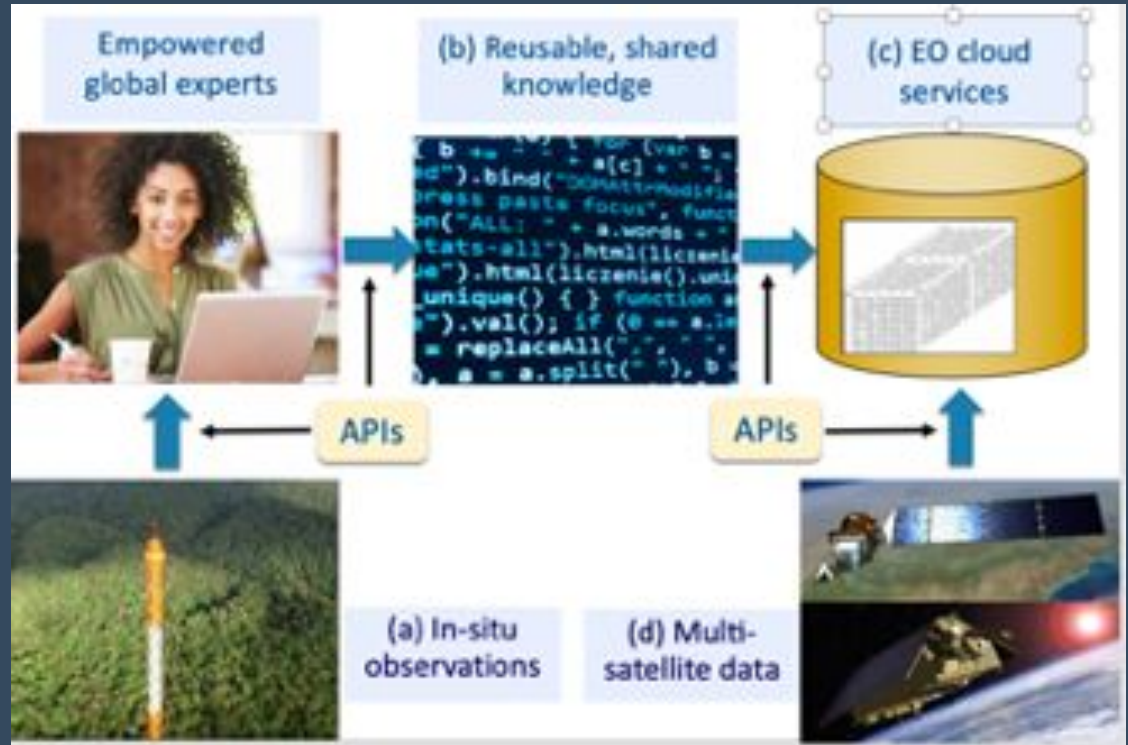


Fig. 1 in GEO EAG Discussion Paper “Building the infrastructure for a results-oriented GEOSS”

GEOSS Solution Space



- JSON-LD*
- schema.org

* Javascript Object Notation - Linked Data



More Federations...

EOSDIS + ...

- ...High-End Computing (Modelers)
- ...Other clouds
- ...End user(s)

Variety-Induced Challenges



- **Recurrent**
 - Semantic divergence (e.g., AM-1, TERRA, Terra, ...)
 - Flexibility vs usability
- **New**
 - Budget control
 - Cost of data movement
- **Meta**
 - Variety of variety
 - Scalability of people
 - Interoperability between concepts (e.g., data ↔ services)