The Value of Being a Trustworthy Repository

Paper IN23A-01, December 10, 2019 Fall AGU Meeting, San Francisco, CA



Jeanne Behnke¹ and Hampapuram (Rama) Ramapriyan^{1,2}

¹ESDIS Project, NASA/Goddard Space Flight Center &

²Science Systems and Applications, Inc.



The Value of Being Trustworthy



- NASA has invested in 60 years of data and information systems
 - But trustworthiness in our repositories is still a goal that we strive to achieve and even exceed
- Why?
 - Enables the success of any NASA science mission
 - Inspires general science research and applications
 - Justifies the cost of operations
 - Contributes to the value of the Open Data Policy
 - Influences the long term, historical view for the data collection
- But what is TRUST and how does it apply to EOSDIS?

EOSDIS Comprises Data of the Whole Earth System

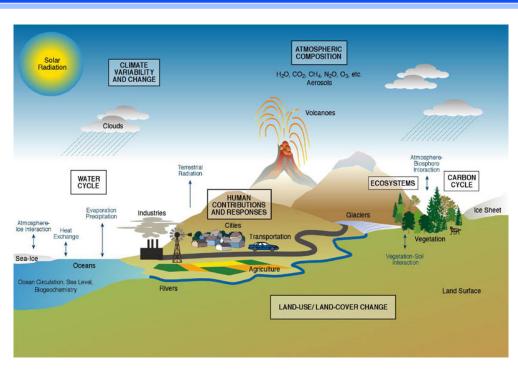


Atmosphere

Winds & Precipitation Aerosols & Clouds Temperature & Humidity Solar radiation Ocean Surface temperature Surface wind fields & Heat flux Surface topography Ocean color Cryosphere Sea/Land Ice Snow Cover Land Cover & Usage Soil Moisture Topography & elevation **Temperature Human Dimensions** Population & Land Use

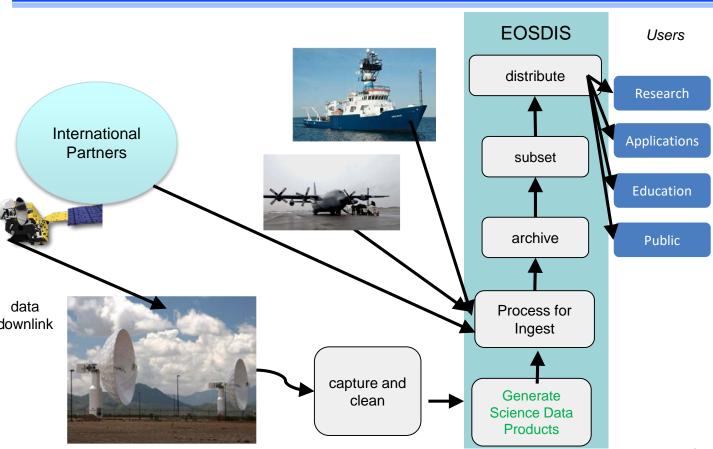
Human &

Environmental Health



Earth Observing System Data and Information System (EOSDIS)





Trustworthy Repositories

Alaska Satellite Facility DAAC

SAR Products, Sea Ice, Polar Processes, Geophysics

National Snow and Ice Data Center DAAC

Frozen Ground, Glaciers, Ice Sheets, Sea Ice, Snow, Soil Moisture

Land Processes DAAC

Land Cover, Surface Reflectance, Radiance, Temperature, Topography, Vegetation Indices

Goddard Earth Sciences Data and Information Services Center

So

Hur

Sy

Global Precipitation, Solar Irradiance, Atmospheric Composition and Dynamics, Global Modeling

Crustal Dynamics Data Information System Space Geodesy

Space Geodesy, Solid Earth

Global Hydrology Resource Center DAAC

Hazardous Weather, Lightning, Tropical Cyclones and Storm-induced Hazards

LaRC Atmospheric Science Data Center

Radiation Budget, Clouds, Aerosols, Tropospheric Chemistry

Distributed Archive

S

sical

rents

aphy DAAC

Ocean Winds.

Circulation &

ea Surface

Laboratory DAAC
Biogeochemical Dynamics,
Ecological Data, Environmental
Processes

Oak Ridge National

TRUST Principles*



- T Transparency is achieved by providing publicly accessible evidence of the services that a repository does and does not offer.
- R Responsibility is a commitment to provide reliable data services.
- U User community is a commitment to implement and enforce the standards and norm of the user community.
- S Sustainability is the capability to support long-term data preservation and use.
- **T Technology** is the infrastructure and capabilities to support the repository operations.

^{*}From: Lin, Dawei, "The TRUST Principles for Trustworthy Data Repositories – An Update, RDA/WDS Repository Certification IG, September 12, 2019

T - Transparency



- All information about policies, practices and capabilities are publicly available
- An extensive, comprehensive and easy-to-traverse website: https://earthdata.nasa.gov/
- Policies (Open data, software and services)
 - https://earthdata.nasa.gov/collaborate/open-data-services-and-software
- > Practices
 - Adding new data https://earthdata.nasa.gov/collaborate/new-missions
 - https://earthdata.nasa.gov/collaborate/new-missions/addingorbital-airborne
 - https://earthdata.nasa.gov/collaborate/new-missions/addingcompetitive-other
 - Requirements, Standards & References -https://earthdata.nasa.gov/esdis/eso/standards-and-references
- Capabilities https://earthdata.nasa.gov/eosdis

R - Responsibility

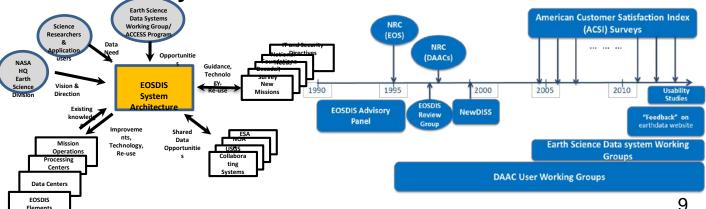


- Commitment to provide reliable data services.
 - Proven record of supporting new missions and data
 - Data products accompanied by information about data quality
 - Sustained funding support for the data collection from NASA
 - Focused attention on discipline-specific user needs; Repositories are collocated with scientific expertise
 - Continuous review of data and metadata collection; migration to new technologies; implementation of new capabilities
 - Commitment to FAIR principles

U - User community



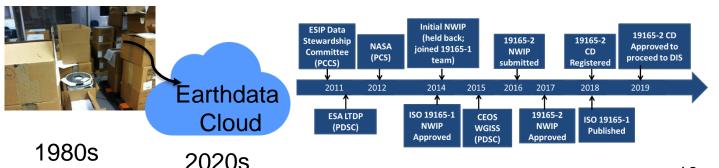
- Commitment to implement and enforce the standards and norm of the user community.
- Proven customer satisfaction as assessed through annual American Consumer Satisfaction Index (ACSI) surveys conducted by the CFI Group
- EOSDIS has 30 years of user community input into design/development
- ESDIS Standards Office reviews and approves community standards for NASA Earth science data



S - Sustainability



- Capability to support long-term data preservation and use.
- Preservation Content Specification to ensure future understandability & reusability
- > Leading ISO standard development
- CoreTrustSeal certification for NASA's EOSDIS repositories located at Centers of expertise
- Sustained long-term commitment and funding from NASA



T - Technology



- Infrastructure and capabilities to support the repository operations.
- On-going evolution since the beginning to keep up with state-ofthe-art and changes in users' expectations
- User needs' assessments, internal/external reviews, Earth Science Data System Working Groups' recommendations help identify capability needs
- Lessons learned and information technology advances coupled with advice/comments from community supports a continuously evolving data system with growing capabilities





Earthdata Cloud

EOSDIS Technology Improvements and System Evolution



Lessons learned and information technology advances coupled with advice/comments from community supports a continuously evolving data system with growing capabilities

Improved

access to

heritage data

Cross-system

order access

interoperability

search and

via data

model

Common

distribution

format (HDF);

other formats

Higher density

offline media

also supported

- Support for high data volume & nearline media
 - Integrated core plus coupled elements
 - Common data model
 - Expanded software tools and services
 - Options to support or interoperate with external data sources

- Coexistence of heterogeneous, distributed data providers / information partners
- Minimal set of core standards; support for communityspecific standards
- Coordinate websites – earthdata.nasa.gov
- Preservation content specifications
- Reusable software
- Service Oriented Architecture
- On-line archives and cross-system service invocation
- Near Real-Time access

- Federated active archives; loosely coupled
- User Needs driven software development
- Common Metadata Repository with a Unified Metadata Model
- Enable user registration to provide persistence and seamless access
- Collaboratively developed data analytics software
- Open Source software and use of GitHub
- Use of commercial cloud resources
- Easy access Browse imagery
- Ease of innovation and technology infusion

Offline media

Discipline/

specific data

mission

systems

specific

proved

only

Community-

standards

■ Data inter-use

cumbersome

Mid-1990s

Late 90s +

2000s

2010s

Conclusion



- As a network of trusted repositories,
 NASA's EOSDIS has supported TRUST principles for over two decades
- Trust is essential for users to have confidence is our data
- Sustenance is essential to trust
- Trust is essential to sustenance

Being a Trustworthy Example



- The reputation of NASA and its Earth science data is golden and is built on the trust of its customers. There are several lessons learned in developing EOSDIS:
 - developing and evolving trustworthy repositories
 - building frameworks for keeping digital collections relevant in an environment that constantly changes
 - being responsiveness to a broad user community with differing perspectives, motives, and cultures
 - being forward-thinking on a flat budget

Thanks to all who make EOSDIS possible

References



- Ramapriyan et al. "Standards and Best Practices Two NASA Examples" IN21D-0876, AGU Fall Meeting, 2019
- Ramapriyan, H. and Behnke, J., 2019. Importance and Incorporation of User Feedback in Earth Science Data Stewardship. *Data Science Journal*, 18(1), p.24. DOI: http://doi.org/10.5334/dsj-2019-024
- Ramapriyan & Lynnes, "Reusability in NASA's Earth Observation System Data and Information System (EOSDIS)", IN 21A-09, AGU Fall Meeting, 2019