Goddard Earth Science Data and Information Center (GES DISC)

http://disc.sci.gsfc.nasa.gov

Steve Kempler
NASA Goddard Space Flight Center

GES DISC Vision

To enable researchers and educators maximize knowledge of the Earth by engaging in understanding their goals, and by leading the advancement of remote sensing information services in response to satisfying their goals

The GES DISC Approach

- Engage Users
  - Communications must be frequent
  - Dedicated points of contact to gather/provide information are identified
- Build economically
  - Look for reuse, ways to save funds
  - Willing to take calculated risks; Otherwise low risk
- But also, build to integrate new technologies
- Engage employees
  - Ensure that employees realize the value of their contributions
  - Treat all employees equally

We will not: Build it and they will come.
We will: Build it because they came (collaborating on mutual interest)

Successes

- Up to date on all Mission reprocessing and documentation
- Successfully released first data products produced for newest missions: GPM, OCO-2, SNPP
- Completed transition to dynamic web page capability, driven by CMR web documentation
- Released new version of Giovanni (smoothing maps, histograms, download Geotiff maps, etc.)
- Delivered the first version of Unified User Interface with support for faceted navigation
- Completed population of HIRDLS data and documentation into preservation system; Other datasets in the works
- Deployed User Registration for GES DISC services and data access
- Recovered heritage data (Nimbus, others) from 1960's vintage media. Rescuing data continues.
- GES DISC is now a recognized data repository by Scientific Data, an open-access, peer-reviewed journal for descriptions of scientifically valuable data sets

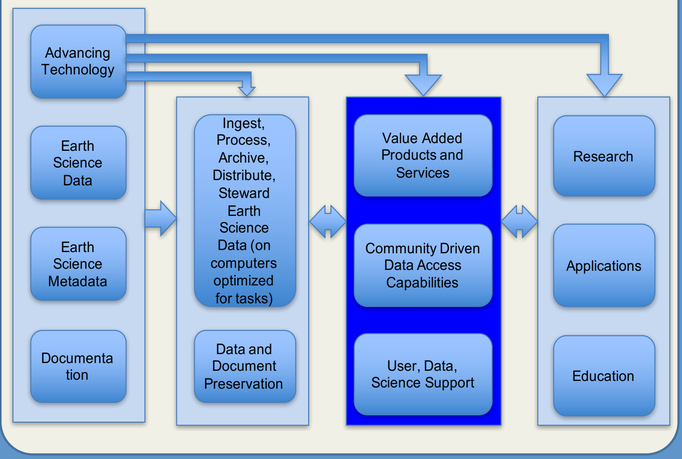
Challenges

- Efficiently adapting potentially useful advancing technologies to specific problems
- Facilitating science research or the masses, based on a handful of use cases
- Rapidly responding to the needs of users needs for value added products, tools, and services
- So much to do… so little time to do it.

Best Practices

- Implement in response to user driven needs
  - UWG, but also user feedback, user surveys, science meetings
- Seek opportunities for collaboration
  - EOSDIS drivers, but also **new** shareable initiatives
- Strategically utilize technology to enhance efficiency in the face of growing archives and number of users
  - Keenly understand forward looking relevant information technologies, but also engage science data users and information technologists
- 'Partner' with users and producers
  - Remain expert in Atmospheric, Hydrology, Climate Modeling data (both NASA and other) and data management services, but also engage science research and applications users to better understand their needs, and improve GES DISC services
- Publish results for the betterment of information science

The GES DISC 'World'



Who We Are: Summary of Expertise

- **Science Data Management** – *Ph.D scientists in applicable earth science disciplines who collaborate with researchers to develop data reduction/exploration tools to facilitate information extraction and multi-mission data coordination; Who understand how data was generated and provide user support*
- **Mission Support** - *Engineers who understand the requirements for costing and sizing information management systems for new or existing missions*
- **Software Engineering** - Engineers who *understand the most effective advanced technologies to further mature data management system usability and efficiency*
- **Operations** - *8 X5 (24 X 7 on call) staff that understands the importance of, and ensures, continuous data ingest, processing, archive and distribution*

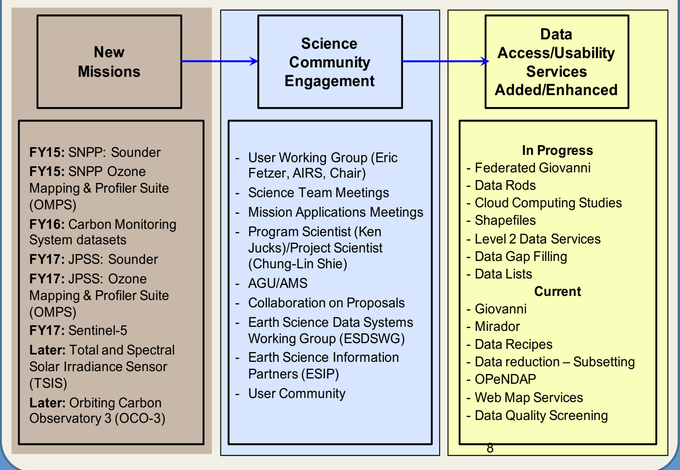
What We Do

- Science User/Data Support
  - Receive and disposition data, science, service inquiries daily
  - Understand and develop new research driven tools and services
  - Analyze metrics to address research need priorities
  - Perform: Outreach, Documentation, Capturing data preservation artifacts
- Mission Support
  - Build tailored archive, distribution, service systems to requirements of new project: Develop ICDs, Ensure formats and metadata guidelines are met, Build hw/sw systems cost effectively to spec, Interfaces work
  - Spearhead Digital Object Identifiers, Landing Pages
- Software Engineering
  - Lead overall system architecture: planning/implementation
  - Implement/maintain flexible system tools and services to enhance data usability, to accommodate evolving user needs
  - Employ advanced S/W Engineering techniques (Agile Methodologies)

What We Do

- Infrastructure
  - Perform System Administration (upgrades, patches, installations, backups, etc.) for main computers and desk/lap-tops
  - Security, web, system configuration management
  - Virtual Machine to support SA
  - Implement and analyze cloud computing application prototyping
- Operations
  - Ensure data ingest, archive, and distribution
  - Apply system monitoring tools to enhance operations efficiency
- Management
  - Manage a diverse staff and set of functions: Contracts, cooperative agreements, budgets, ~60 people, new business, system ownership, etc.

Stewarding Mission Data



Goddard Earth Sciences Data and Information Services Center (GES DISC)

Atmospheric Composition
- Total Ozone Mapping Spectrometer (TOMS)
- Upper Atmosphere Research Satellite (UARS)
- Solar Radiation and Climate Experiment (SORCE)
- Aura: Ozone Monitoring Instrument (OMI), High Resolution Dynamics Infrared Sounder (HIRDLS), Microwave Limb Sounder (MLS)
- Atmospheric CO2 Observations from Space (ACOS)
- Historical datasets from Nimbus, Tiros, SME, others
- Orbiting Carbon Observatory 2 (OCO-2)
- SNPP: Ozone Mapping & Profiler Suite (OMPS)
- FY16: Carbon Monitoring System (CMS)
- FY17: JPSS: Ozone Mapping & Profiler Suite (OMPS)
- FY17: Sentinel 5
- Later: Total and Spectral Solar Irradiance Sensor (TSIS)
- Later: Orbiting Carbon Observatory 3 (OCO-3)

Modeling
- Modern-Era Retrospective Analysis For Research and Applications (MERRA)
- Global Land Data Assimilation System (GLDAS)
- North American Land Data Assimilation System (NLDAS)

Atmospheric Dynamics
- TIROS Operational Vertical Sounder (TOVS)
- Aqua: Atmospheric Infrared Sounder (AIRS)
- SNPP: Sounder
- FY17: JPSS: Sounder

MEaSUREs Data
- MEaSUREs 2006
- MEaSUREs 2012

Precipitation
- Tropical Rainfall Measuring Mission (TRMM)
- Hydrology Collections
- Global Precipitation Mission (GPM)

Current Operational Services/Tools

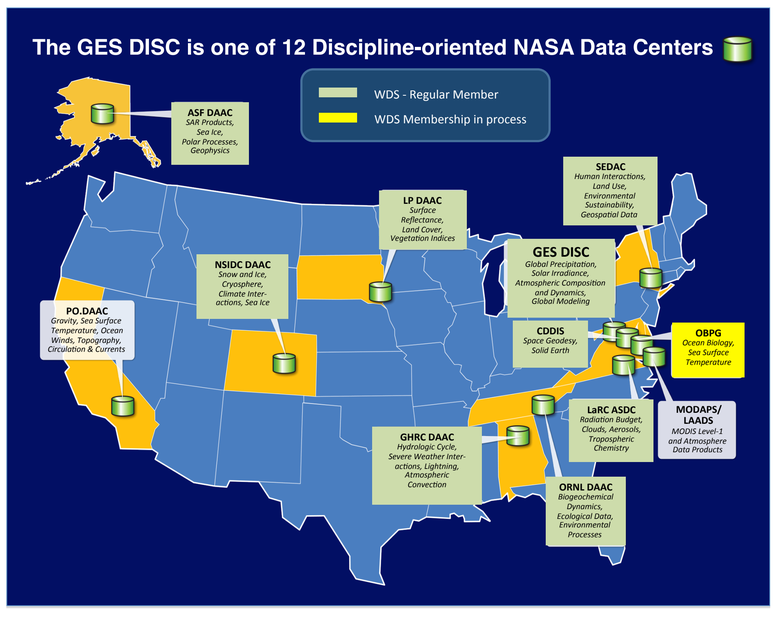
- **Giovanni** – Data Discovery, Visualization and Exploration
- **Mirador** – Data Search and access
- **Simple Subset Wizard** – Cross DAAC effort to provide subsetting
- **Data Recipes**
- **OpenDAP & GrADS Data Server**
- **Open Geospatial Consortium (OGC) Web Map Service (WMS)**
- **Data provided in various formats** (HDF, netCDF. ASCII, kmz, others)
- **Shapefiles** – Significant for applications
- **Data Reduction** – Subset desired measurements from data products
- **Applications Contributions**: Applied Remote Sensing Training (ARSET); Hydrologists (CUAHSI); USDA World Board; Public Health
- **User Registration**
- **Digital Object Identifier and Landing Pages**
- **Data Stewardship**

Coming:
- **Cloud applications**
- **Unified User Interface**

Leadership Activities

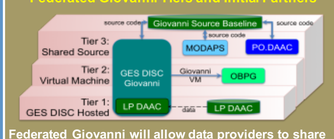
ESDSWG Participation
- Leader:
  - Virtual Collections (Completed)
  - Time series – (Co-lead)
  - Data Interoperability (Co-lead)
  - Level 2 Data Visualization (New)
  - Atmospheric Science User Forum (Co-lead, New)
- Participant:
  - Data Recipe (Completed)
  - Data Quality, Geospatial Web Services Best Practices, OPeNDAP, Search Relevance

ESIP Participation
- Leader:
  - Earth Science Data Analytics
  - Agriculture and Climate
- Participant:
  - Education, Preservation and Stewardship, OpenSearch, Data Quality



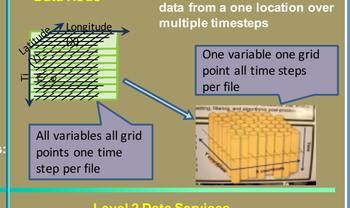
GES DISC: Looking to the Future

Federated Giovanni Tiers and Initial Partners



Federated Giovanni will allow data providers to share their data utilizing Giovanni services in either of 3 ways: GES DISC hosted, Virtual server machine, Hosted on providers machine

Data Rods



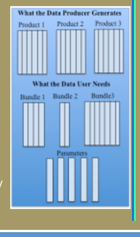
Makes it possible to retrieve data from a one location over multiple timesteps

DATA LISTS

Requests from users:
- Data bundled together for multi-measurement analysis
- Variables, not products containing a lot of variables

Data Lists:
- Break out products into variables
- Bundle data locations by science themes, events, geographic location, etc.

Impacts: quicker access, less work by scientist



Level 2 Data Services

Data & Quality in Level 2 Visualization (DQL2Viz)

GES DISC Facts – as of 2015

Archive Volume (as of 7/16):	1,441 TB
Distribution Volume:	2,071 TB (up 100% in 4 years)
Granules Distributed:	5,427 M (up ~100% in 3 years)
NRT Distribution Volume:	38.7 TB (up 150% in 3 years)
NRT Granules Distributed:	8,707 K (up ~100% in 2 years)
Number of Unique Users:	92,341 (almost 100% in 4 years)
GES DISC Presentations/Publications:	~25/year
Number of Publications citing Giovanni:	>200/year
Number of Recipes (currently):	31
Number of DOIs registered (currently):	384
Data Collections with Landing Pages:	All

NRT = Near Real Time

WDS Contacts: Steven.J.Kempler@nasa.gov