Role and Evolution of NASA's Earth Science Data Systems

IEEE EDS/CAS Chapter –
Camarillo, CA
August 19,2015

H. K. Ramapriyan
Science Systems and Applications, Inc./
NASA GSFC ESDIS Project
Hampapuram.Ramapriyan@ssaihq.com

Acknowledgement

- This presentation is the result of my long association with the NASA Earth Science Data and Information System Project
- It has evolved over the years with inputs from several members of the Project and has been used in various forms in many presentations
- My present work with the ESDIS Project is supported under contract number NNG15HQ01C with NASA Goddard Space Flight Center

Topics

- NASA's Earth Science Data Systems
  - Core and Community Capabilities
- Earth Observing System Data and Information System (EOSDIS)
  - Scope and Context
  - System of systems
- Recent Developments
- EOSDIS Evolution - Community inputs
- IEEE GRSS Earth Science Informatics Technical Committee and EOSDIS
- Conclusion

NASA's Earth Science Data Systems

- **"Advance understanding of Earth and develop technologies to improve the quality of life on our home planet." -- *2014 NASA Strategic Plan***
- **NASA's Earth Science Data Systems directly support this strategic goal by providing end-to-end capabilities to deliver data and information products to users**
- **NASA's Earth Science Data and Information Policy promotes usage of data by the community**
  - **No period of exclusive access - Data are available after initial checkout**
  - **Data available at no cost to all users on a non-discriminatory basis except where agreed upon with international partners**

Core and Community Capabilities

- **Core Capabilities**
  - **Basic operational capabilities to process, archive, manage and distribute data from NASA missions**
    - **Earth Observing System Data and Information System (EOSDIS)**
    - **Precipitation Processing System – NASA GSFC**
    - **Laboratory for Atmospheric and Space Physics (LASP) Interactive Solar Irradiance Data Center - University of Colorado**
    - **CloudSat Data Processing Center – Colorado State University**
- **Community Capabilities**
  - **Peer-review-selected projects**
  - **New data products – Making Earth System Data Records for Use in Research Environments (MEaSUREs)**
  - **Research in Earth Science Informatics to feed into the evolution of the core components**
    - **Applied Information Systems Technology (AIST)**
    - **Advancing Collaborative Connections for Earth System Science (ACCESS)**

EOSDIS

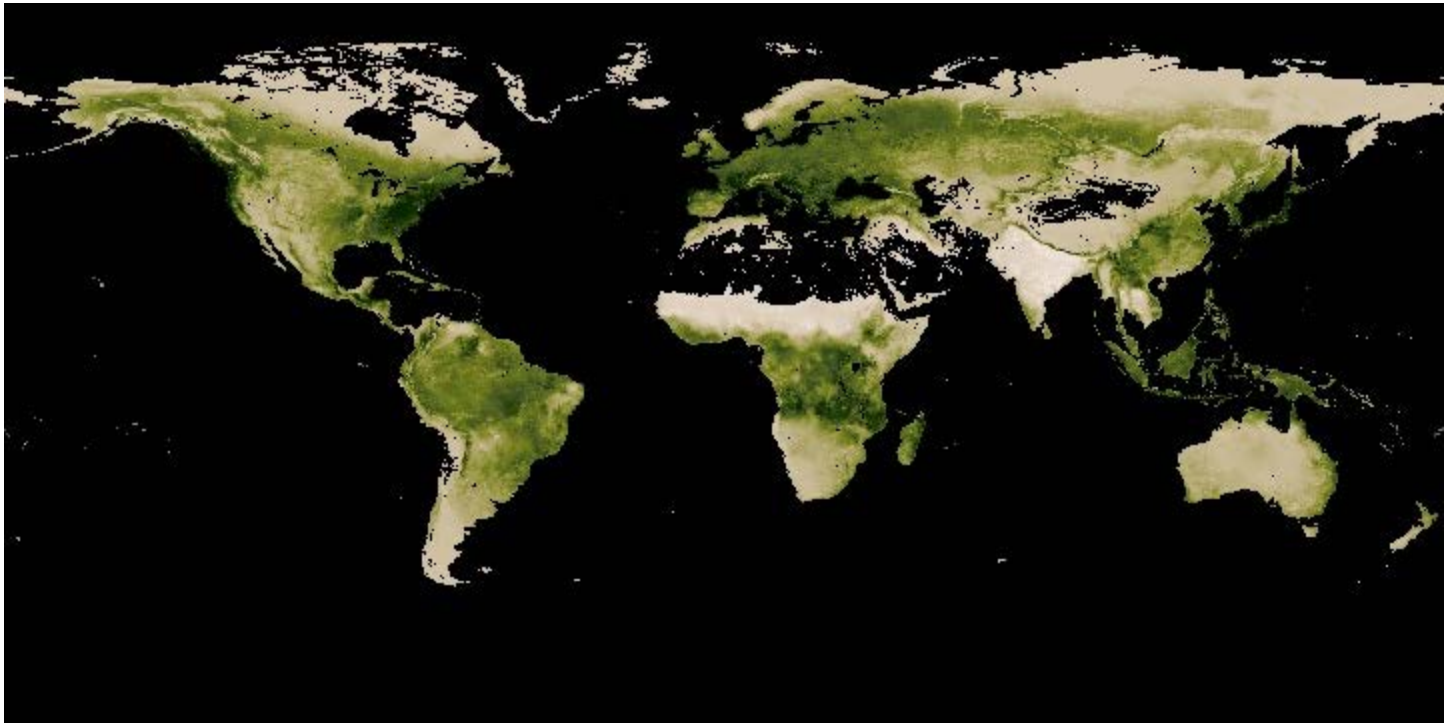
- **Major core capability in NASA's Earth Science Data Systems Program.**
- **Provides end-to-end capabilities for managing NASA's Earth science data.**
  - **Science Operations**
    - **Science data processing**
    - **Data management**
    - **Interoperable distributed data archives**
    - **On-line data access services**
    - **Earth science discipline-oriented user services**
  - **Network Data Transport to distributed system elements**

Extensive Data Collection

- **> 8200 data types (collections)**
  - Land
    - » Cover & Usage
    - » Surface temperature
    - » Soil moisture
    - » Surface topography
  - 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

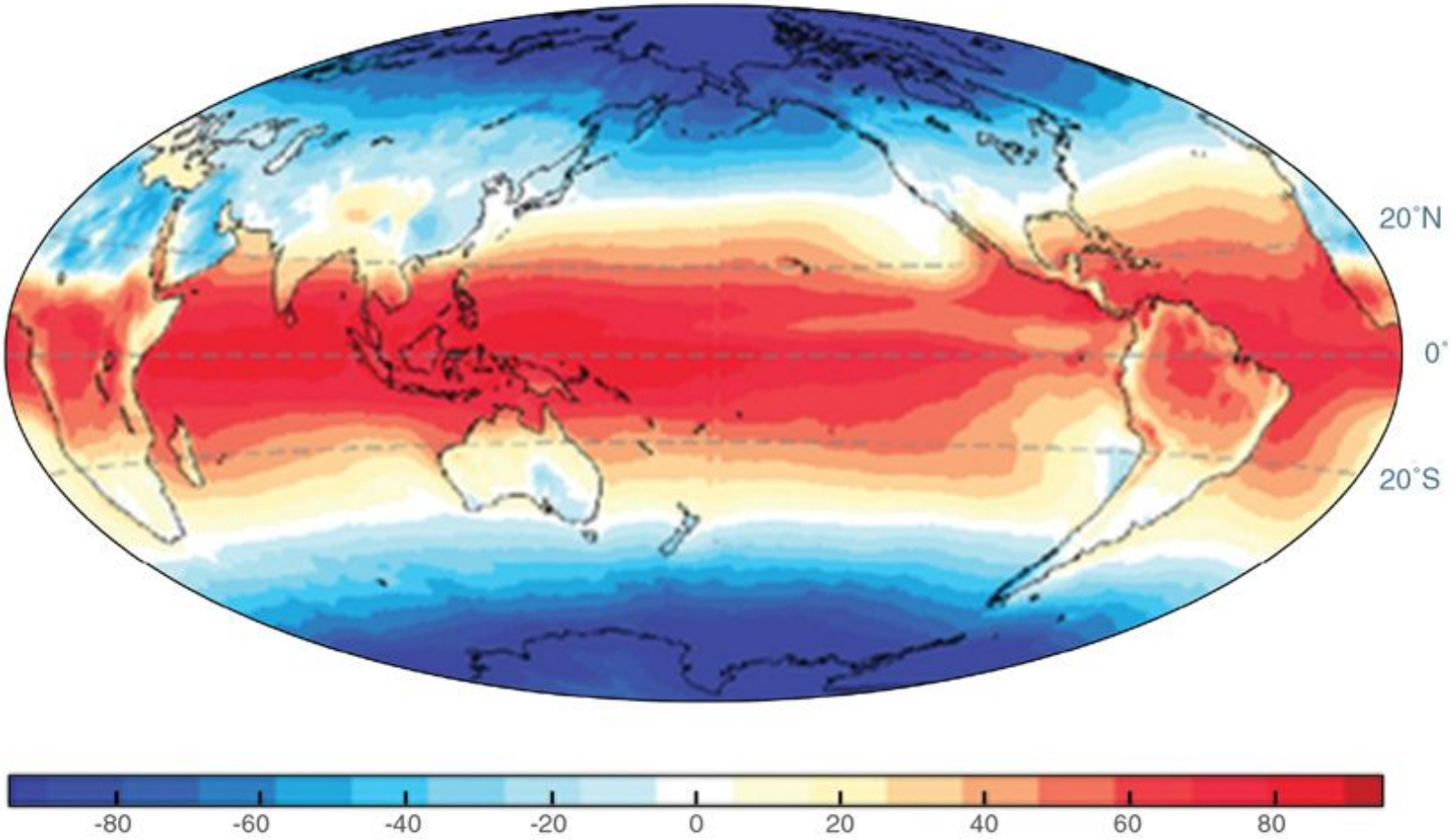
- Human Dimensions
  - » Population & Land Use
  - » Human & Environmental Health
  - » Ecosystems

Global Net Primary Productivity



Net Primary Productivity is the amount of carbon absorbed by plants minus carbon released by plants, measured in grams of carbon per square meter per day. Image shows the averages over May 15, globally. Credits - Image made by Reto Stockli, NASA's Earth Observatory Team, using data provided by the MODIS Land Science Team

Top of Atmosphere Radiation



At the top of the atmosphere (TOA), incoming and outgoing radiation determine Earth's average temperature. This image shows averaged net downward TOA radiation from the Clouds and Earth's Radiant Energy System (CERES) instrument from 2001 to 2010. The Southern Hemisphere receives more net radiation than the Northern Hemisphere. (Courtesy D. Frierson et al., 2013, Nature Geoscience) – accessed through https://earthdata.nasa.gov/user-resources/sensing-our-planet/rooting-out-rainfall

Air Quality in Northeastern China (1 of 2)

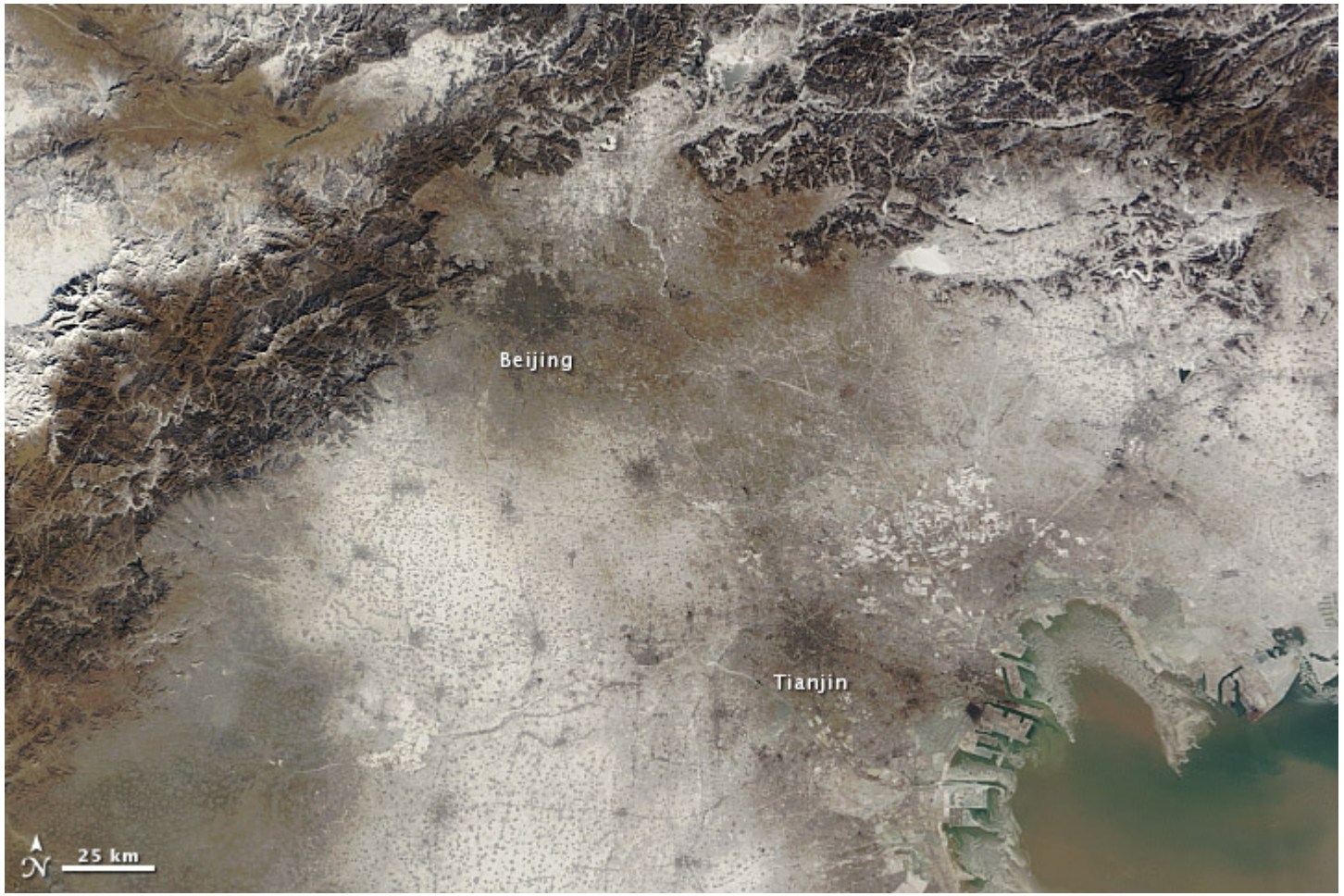


Image acquired by Moderate Resolution Imaging Spectroradiometer (MODIS) on NASA's Terra satellite on January 3, 2013 - NASA image courtesy Jeff Schmaltz, LANCE MODIS Rapid Response.

Air Quality in Northeastern China (2 of 2)

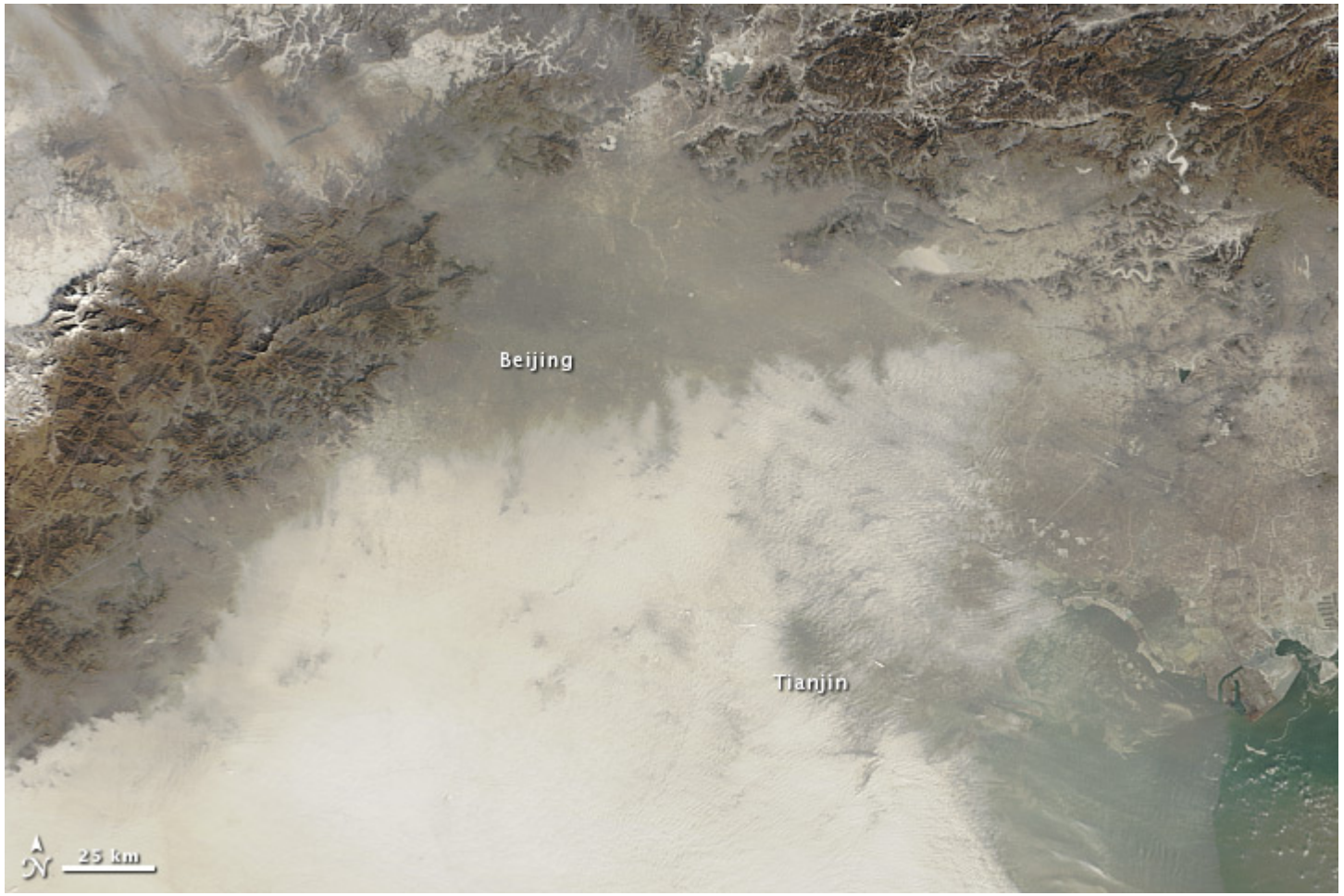
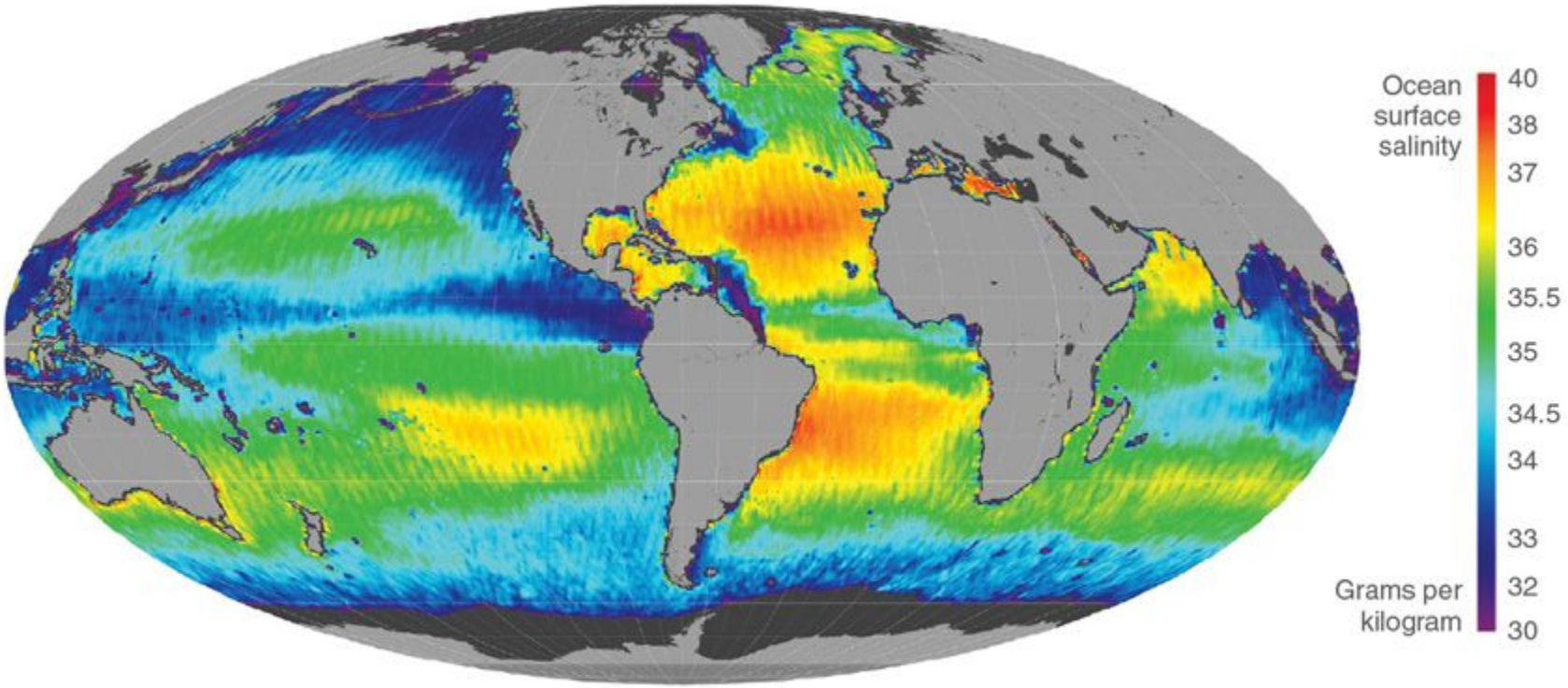


Image acquired by Moderate Resolution Imaging Spectroradiometer (MODIS) on NASA's Terra satellite on January 14, 2013 - NASA image courtesy Jeff Schmaltz, LANCE MODIS Rapid Response.

Sea Surface Salinity



This image of Aquarius sea surface salinity (SSS) measurements averaged for 2012 shows a global color scale of salinity intensity. Warm colors mark stronger salinity values. Values are shown in a range between 30 grams per kilogram (purple) and 40 grams per kilogram (red). (Courtesy N. Kuring/NASA) – accessed through https://earthdata.nasa.gov/user-resources/sensing-our-planet/salt-of-the-sea.

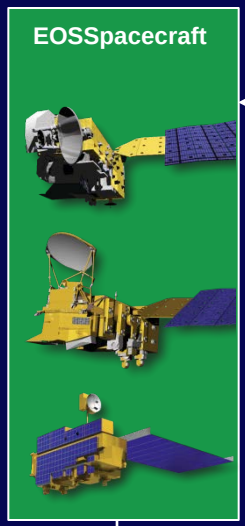
National Aeronautics and Space Administration

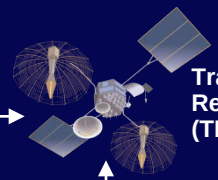
Earth Science Data Operations

Mission Operations

Science Operations

Data Acquisition	Flight Operations, Data Capture, Initial Processing, Backup Archive	Data Transport to Data Centers/ SIPSs	Science Data Processing, Data Management, Interoperable Data Archive, and Distribution	Distribution and Data Access




Tracking and Data Relay Satellite (TDRS)


White Sands Complex (WSC)


EOS Data Operations System (EDOS) Data Processing


EOS Polar Ground Stations


EOS Operations Center (EOC) Mission Control

Direct Broadcast (DB)


Direct Broadcast/ Direct Readout Stations

NASA Integrated Services Network (NISN) Mission Services


EOSDIS Sci. Data Centers


Instrument Teams and Science Investigator-led Processing Systems (SIPSs)

Internet (Search, Order, Distribution)

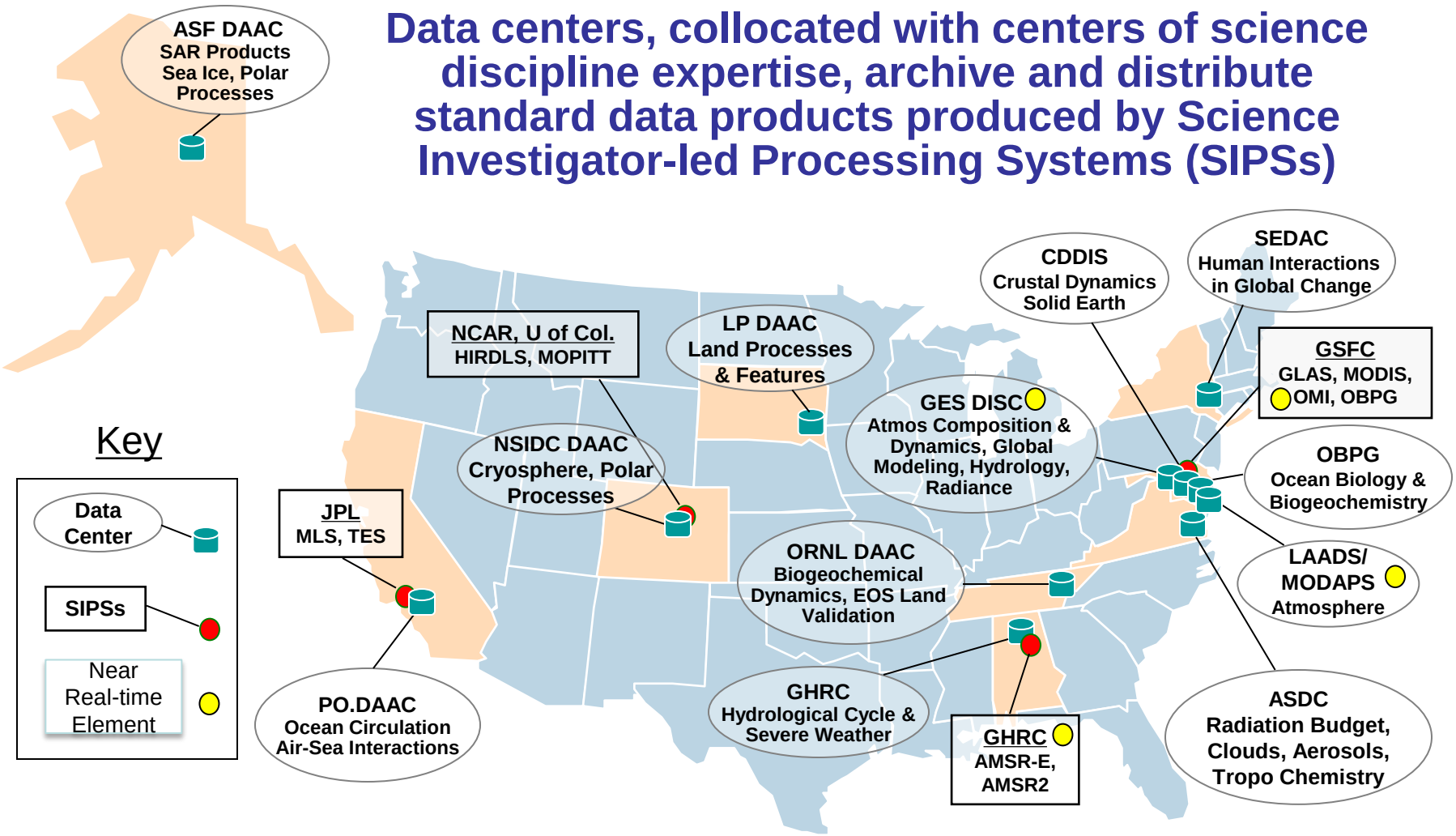
- Research
- Education
- Value-Added Providers
- Interagency Data Centers
- Earth System Models
- International Partners
- Decision Support Systems

EOSDIS – A System of Systems

- **Instrument and science expertise needed to process data**
  - **Most EOS standard products are generated at Science Investigator-led Processing Systems (SIPSs) under supervision of PIs**
- **Earth Science discipline knowledge needed to ensure data stewardship**
  - **Processed data are archived and distributed by discipline-specialized EOSDIS Science Data Centers (Distributed Active Archive Centers – DAACs)**
- **Expertise in system interoperability needed to provide cross-system (interdisciplinary) data access**
  - **EOS Clearing House (ECHO) middleware and associated clients provide search and access to data across all EOSDIS Science Data Centers**
- **EOSDIS data collections are diverse:**
  - **Primary sources are instruments on-board NASA spacecraft**
  - **Ancillary, airborne, in-situ and socio-economic data**
  - **Data from international partners**
  - **Comprehensive approach to multi-discipline science**
  - **Feed growing need by models (e.g., climate models)**

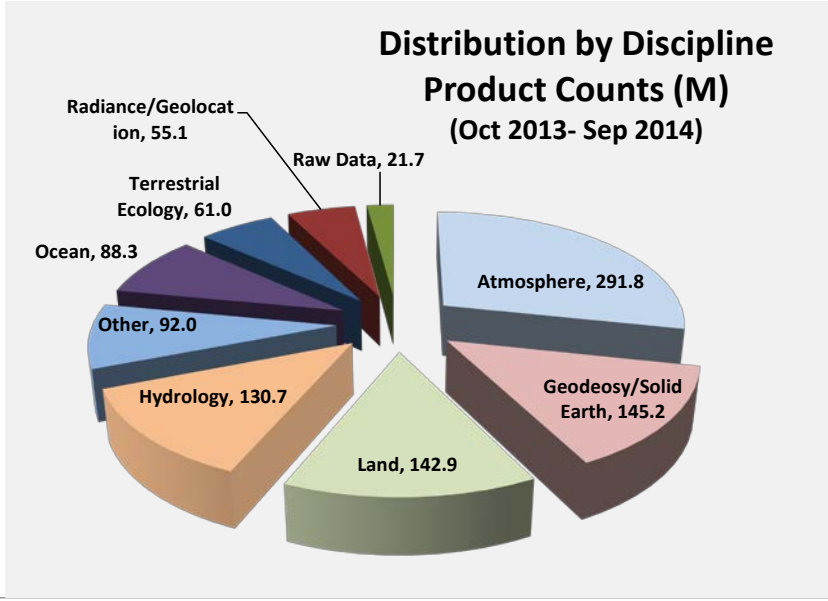
EOSDIS Facilities

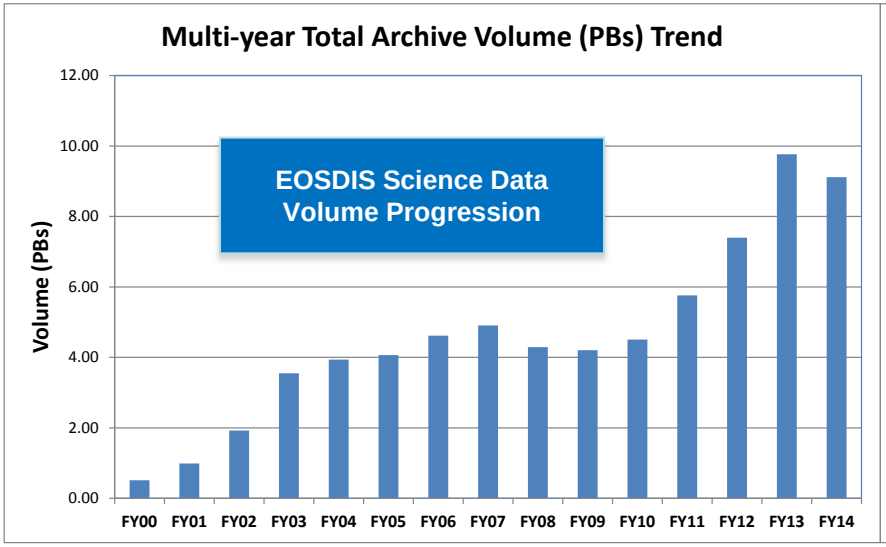
Data centers, collocated with centers of science discipline expertise, archive and distribute standard data products produced by Science Investigator-led Processing Systems (SIPSs)

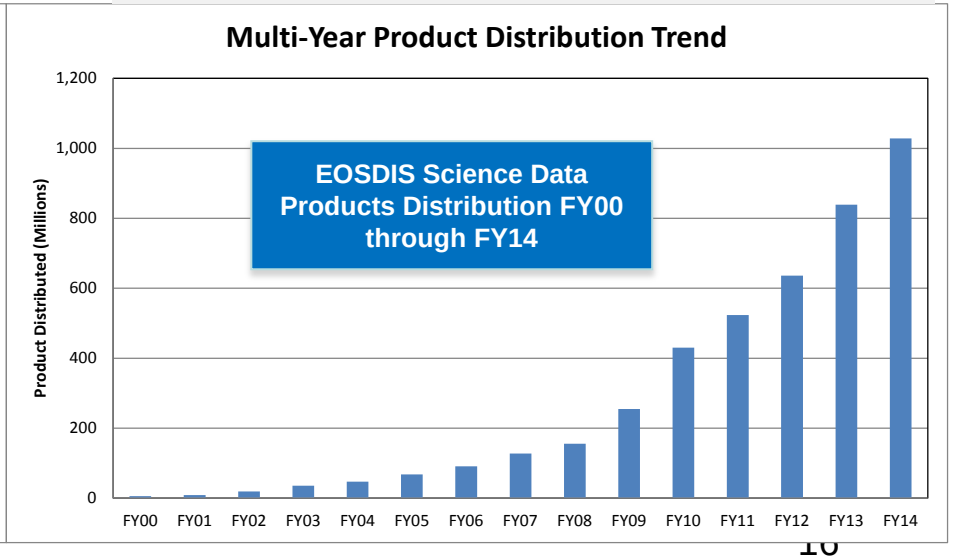


EOSDIS Key Metrics

EOSDIS FY2014 Metrics (Oct. 1, 2013 to Sept. 30, 2014)	
Unique Data Sets	8,292
Distinct Users of EOSDIS Data and Services	2.0 M
Web Site Visits	2.3 M
Average Archive Growth	6.4 TB/day
Total Archive Volume	9.1 PB
End User Distribution Products	1,028 M
End User Average Distribution Volume	27.9 TB/day







Recent and On-Going Developments (1 of 2)

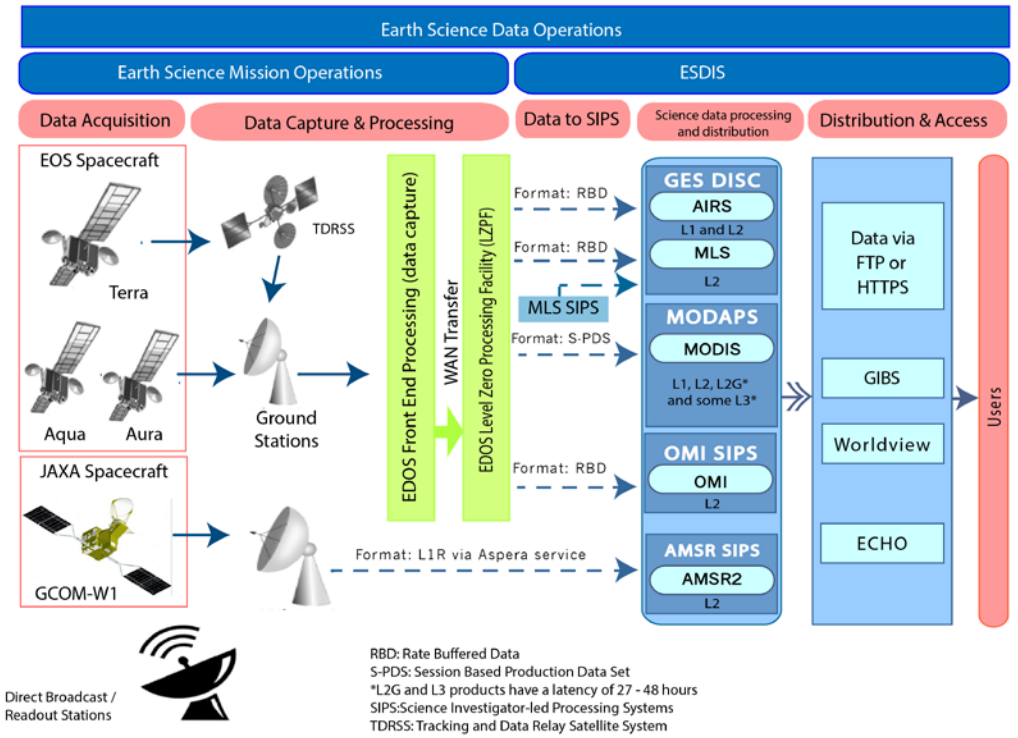
- **Land and Atmosphere Near real-time Capability for EOS (LANCE)**
- **Coherent Web Interface: http://earthdata.nasa.gov is operational**
  - **Provides a unified view of NASA Earth science data system resources**
  - **Consolidates 14 web sites, and provides links to various ways to access data and to related external sites**
- **User Registration System – uniform approach to registration across EOSDIS components**
- **Global Imagery Browse Services (GIBS)**
  - **Standards-based, full resolution, interactive browse capability**
  - **Accessible from http://earthdata.nasa.gov wiki**

Recent and On-Going Developments (2 of 2)

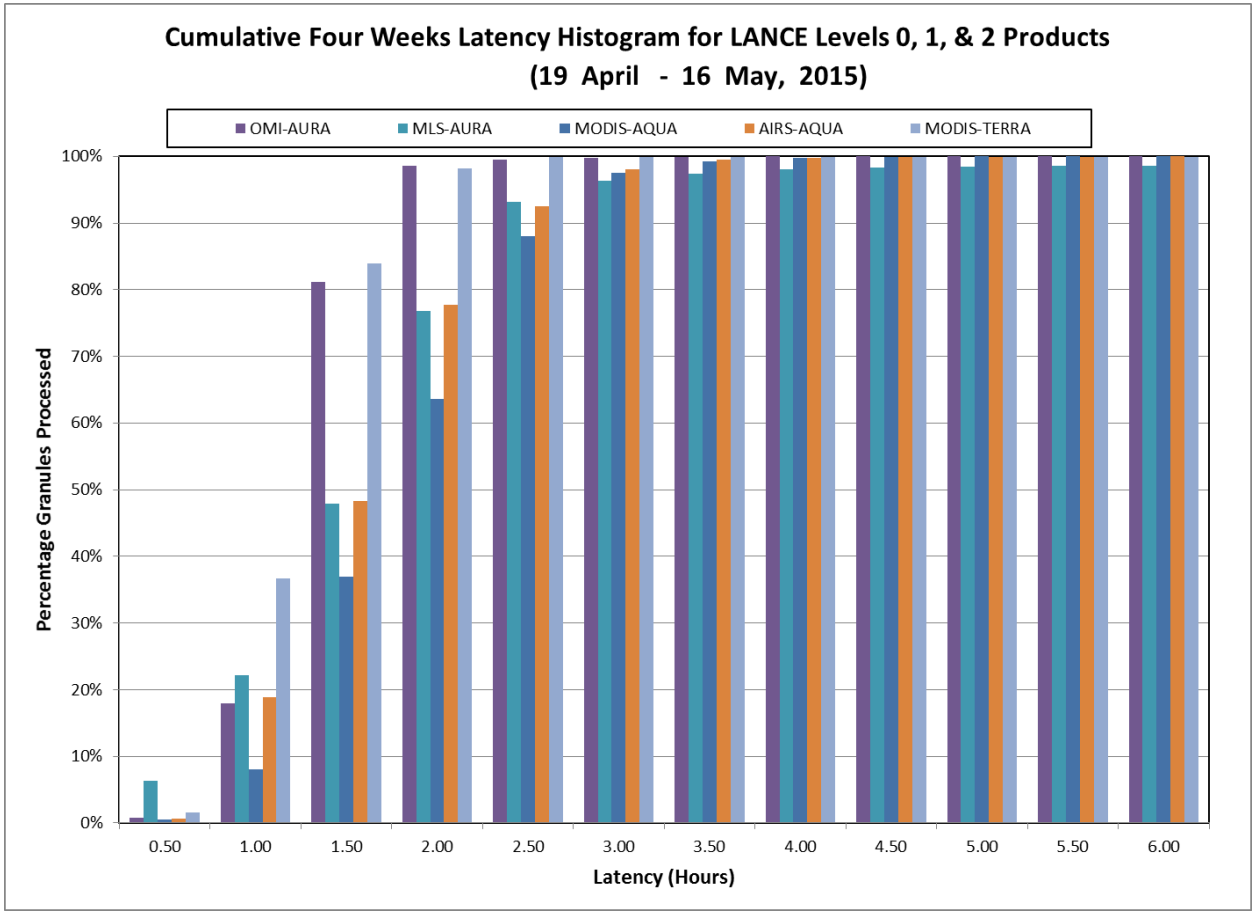
- **Metadata Architecture Study**
  - Initial Study made recommendations on adopting a common approach to metadata to improve user experience and reduce efforts by data providers
  - Phased approach to implementing recommendations
- **Unified Metadata Model and Common Metadata Repository**
- **Big Earth Data Initiative (BEDI)**
- **Preservation Content Specification**
- **Digital Object Identifiers**
  - ESDIS Project is a registration authority (prefix 10.5067)
  - DOI assignments to datasets in progress

Land, Atmosphere Near-real-time Capability for EOS (LANCE)

- **Building on existing EOSDIS elements provides data from MODIS, OMI, AIRS, MLS, and AMSR instruments in near real-time (< 3 hours from observation)**
- **Utilizes software for Standard Science Products, but relaxes requirements for ancillary data inputs**
- **High operational availability**
- **Applications of LANCE data include:**
  - **Numerical weather & climate prediction/forecasting**
  - **Monitoring of Natural Hazards**
  - **Disaster Relief**
  - **Agriculture**
  - **Air quality**
  - **Homeland Security**



LANCE Latencies



Over the four weeks indicated above, >97% of near real-time data requests were satisfied within 3 hours.

EOSDIS Evolution: Earthdata Website

- What is the Earthdata Website?
  - Earthdata was created as a sustainable, evolvable, and reliable Website that represents our community's needs for NASA Earth science data and information.
  - It was designed to support collaboration within and between organizations, and for development and integration of new applications.
  - It addresses the need for a coherent and comprehensive Web presence of the Earth Science Data Systems Program.
  - See Earthdata at https://earthdata.nasa.gov/.

- Benefits of the Earthdata Website:
  - Better represents EOSDIS programmatic investments and capabilities.
  - Presents data centers more clearly as elements within a larger system of systems.
  - Facilitates multidisciplinary research and data integration.
  - More quickly responds to emerging technologies
  - Provides a platform for demonstration of interoperability throughout all of our systems.

EOSDIS Evolution: Worldview and Global Browse Imagery Services

GIBS / Worldview Goal:

To transform how users interact with and discover NASA Earth data; make it visual

Approach:

- The **Global Imagery Browse Services (GIBS)** provide open access to full resolution imagery derived from NASA products to any mapping client and script
  https://earthdata.nasa.gov/gibs

  Open-Access Servers

- **Worldview** is an open source, browser-based client to interactively explore GIBS (and SEDAC) imagery and download the underlying data
  https://worldview.earthdata.nasa.gov

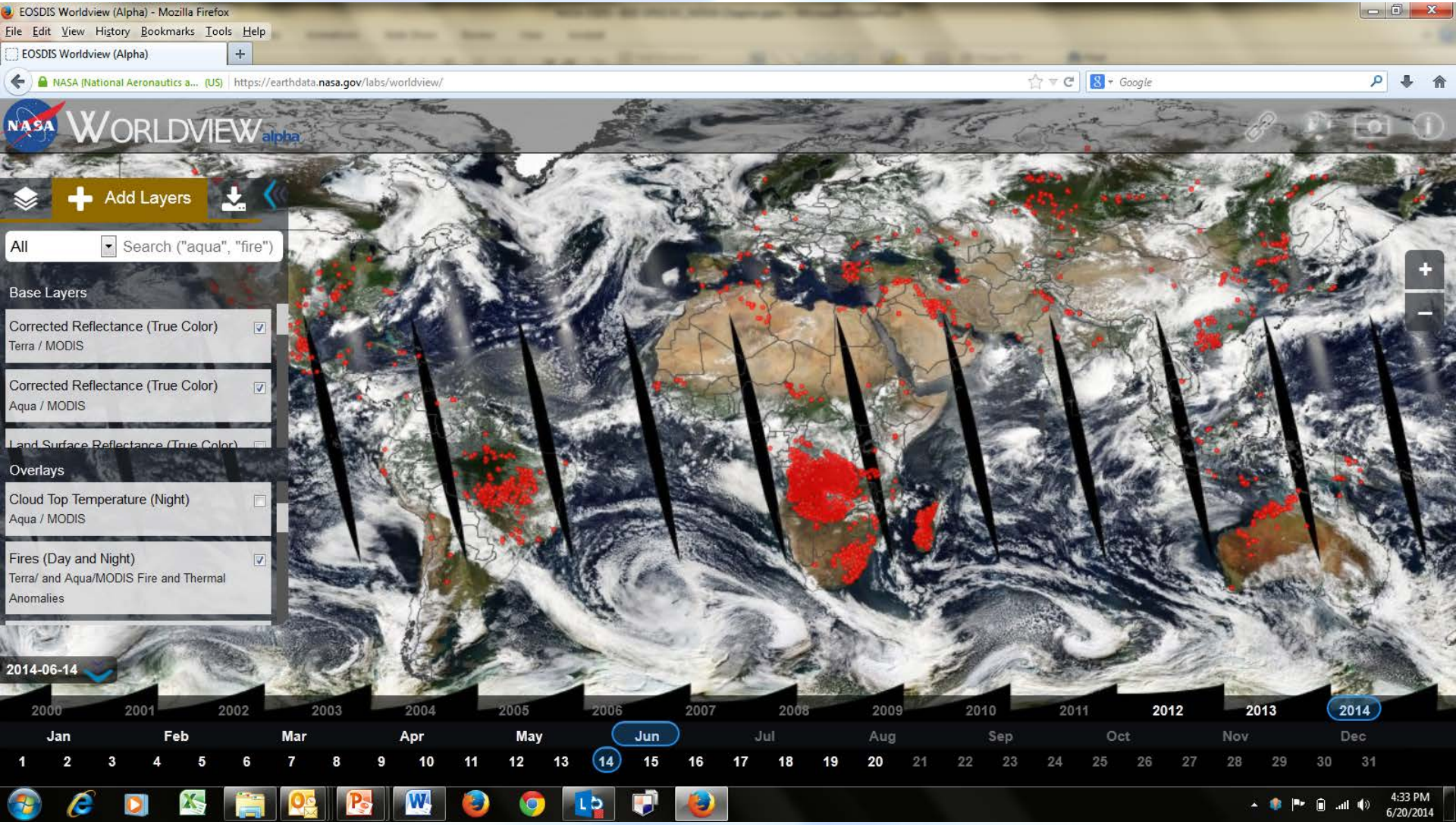
  Client

Global Image Browse Service (GIBS)

- Goal: “Parameter Visualizations” for all EOSDIS Imagery
- Standardized access via OGC WMTS / TWMS / WMS / KML
- Source code for the GIBS OnEarth server and sample code available at the GIBS GitHub site
- Repository of pre-prepared, hierarchically stored imagery to maximize performance for “full-resolution” browse
- Clients can be built to use and display images in GIBS – WorldView is an example

Worldview: Reference Client for GIBS

http://earthdata.nasa.gov/worldview
http://earthdata.nasa.gov/gibs



EOSDIS Evolution - Community Inputs

- **EOSDIS Advisory Panel – aka Data Panel (early to mid-1990's)**
  - **adhere to a flexible, distributed, portable, evolutionary design;**
  - **distribute data products by appropriate high-bandwidth communication or other media;**
  - **operate prototypes in a changing experimental environment**

  →Distributed architecture with DAACs

  →Version 0 working prototype
- **NRC Review (1995)**
  - **"Responsibility for product generation and publication and for user services should be transferred to a federation of partners selected through a competitive process open to all"**

  →Working Prototype Earth Science Information Partners (ESIP) Federation

EOSDIS Evolution - Community Inputs

- **EOSDIS Review Group (1997)**
  - **Recommended "an adaptive approach which will be less centralized, giving more responsibility to the PIs"**

  → Science Investigator-led Processing System (SIPSs)
- **New Data and Information Systems and Services (NewDISS) Strategy Team (1998 - 2002)**
  - **Six recommendations**
    - **Clearly define components**
    - **Employ Infrastructure providing NASA-private sector liaisons**
    - **Employ competitive processes to select components**
    - **Empower science investigators for data system development, processing archiving and distribution**
    - **Apply lessons learned from WP-ESIP Federation**
    - **Charter transition team**

  → Core and Community Data Systems (Core: EOSDIS with DAACs; Community: REASoN projects → ACCESS & MEaSUREs)

  → ESIP Federation

  → Strategic Evolution of Earth Science Enterprise (ESE) Data Systems (SEEDS) Study → Earth Science Data System Working Groups (ESDSWG, 2004)

EOSDIS Evolution - Community Inputs

- **Evolution of EOSDIS Elements Study Team/ Technical Team (2005)**
  - **Developed "EOSDIS 2015 Vision"**

  →First step implementation during 2006-2008 – reallocated functions, simplified system, increased automation, improved services, reduced operations costs

  →Vision tenets continue to be used as a checklist to assess progress of on-going improvements
- **DAAC User Working Groups (on-going)**
- **American Customer Satisfaction Index Surveys (annual)**

EOSDIS Evolution - Community Inputs

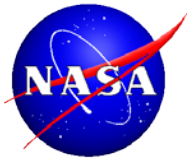
- **Earth Science Data System Working Groups (on-going)**
  - **focus on exploration and development of recommendations derived from pertinent community insights**
  - **organized around key technology and information system issues**
  - **Members from ACCESS, MEaSUREs, DAACs, ESDIS**
  - **2014-2015 Working Groups**

❑ Airborne Data
❑ ASCII for Science Data
❑ Cloud Computing
❑ Data-Intensive Architecture
❑ Data Preservation Practices
❑ Data Quality
❑ Data Recipes
❑ Dataset Interoperability
❑ Digital Object Identifiers
❑ Geospatial
❑ Innovations Lab
❑ Open Source
❑ Provenance for Earth Science (PROV-ES)
❑ Technology Infusion
❑ Vision 2020
❑ Visualization

ESDSWG ←→ IEEE GRSS ESI TC

	ESDSWG	Earth Science Informatics Technical Committee
Processing	• Cloud Computing • Data-Intensive Architectures	• Cloud Computing • Spatial/Temporal analysis Tools • Earth system modeling tools
Archiving/ Stewardship	• Data Preservation Practices • Data Quality • Digital Object Identifiers • PROV-ES	• Preservation • Quality • Data stewardship • Provenance
Access	• Recipes • Dataset Interoperability • Visualization • Geospatial	• Knowledge representation and information models • Cyberinfrastructures • Interoperability and standardization • Data discovery and access • Web-based services and analysis • Geospatial information, knowledge, and decision support systems
Evolution/ Technology	• Innovations Lab • Open Source • Technology Infusion • Vision 2020	• Emerging information technologies and their applications in the geosciences • Sensor web and applications • spatial and process ontologies, vocabularies • semantic web
Other	• Airborne Data • ASCII for Science Data	• Data and information policies

Vision 2020



Discovery and Access

- **Machine Level Discovery and Access** for all data.
- **Seamless Cross-agency Discovery**.
- **Dataset Selection Guidance** based on fitness for purpose.
- **Metadata Naming Conventions** for Variables, Platforms, Instruments, Resolution…
- **Virtual Collections** oriented around science problems.

Integration

- **NASA data can be combined** with data from other agencies and nations.
- **Combining Tools and Services** within the community is easy.
- **Enable sharing** of any scientific resource.

Usage

- **Intelligent Tool Catalogs** suggest tools to work with the data.
- **Publications are linked** to data and tools that allow interactions with the data.
- **Automatic Mobile Data and Processing** to achieve optimal performance.
- **Quantitative Quality** for all data.
- **Reproducibility** of research results with high precision.
- **Documentation** is Concise, Comprehensive and Consistent for all data variables.
- **Capacity Building** mechanisms for people with limited literacy in science, technology, and/or English.
- **Data Analysis at Scale** over any arbitrarily defined area.
- **Dataset Upgrading** for high-value datasets to fully support rich capabilities.

DRAFT

Common Metadata Repository (CMR)

Goals

- Designed to improve the discovery and access of NASA data, CMR will provide a single source of unified, high-quality, and reliable Earth Science metadata while merging the inventories of ECHO (EOS Clearing House) and GCMD (Global Change Master Directory) / IDN (International Directory Network). CMR will be the authoritative management system for all metadata of EOSDIS data holdings.
- CMR also provides a metadata model that documents vital elements that may be represented across various metadata formats and standards and unify them through core fields useful for data discovery and service invocations.

Current Status

- By the end of CY 2015, CMR will have released the minimal set of functionality that addresses the major goals laid out in a CMR Operations Concept including, sub-second search response, unification of ECHO and GCMD/IDN metadata, enhanced metadata management capabilities, and a unified (and expandable) metadata model representing collections, granules, and services.

Next Steps

- In 2016, CMR will begin to implement enhancements to quality assessment and assurance, search relevancy ranking, science keyword support and ontology and service initiation and chaining.

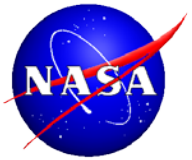
Big Earth Data Initiative (BEDI)

- **Background:**
  - **In 2013, the White House Office of Science and Technology Policy (OSTP) kicked off the Big Earth Data Initiative (BEDI) as a multi-agency (NASA, NOAA, USGS) effort to make the collection of Earth Observation (EO) data more readily available and useful to users.**
  - **The data directly supports 12 Societal Benefit Areas (SBAs).**
  - **Funding provided to NASA FY14, FY15**
- **BEDI Objectives:**
  - **Discovery – Make finding of datasets *simpler***
  - **Accessibility - Make datasets readily *available* to users**
  - **Usability – Provide services to *use* datasets**
- **Task objectives for EOSDIS DAACs:**
  - **Provide metadata to EOSDIS Common Metadata Repository (CMR)**
  - **Maintain persistent identifiers for data collections via Digital Object Identifiers (DOIs)**
  - **Make data available online via OPeNDAP or some other useful service**
  - **Make imagery available in Worldview via GIBS**
  - **Report metrics**

Preservation

- NASA is not a “permanent archive” agency
  - Must maintain “research archive” for as long as data are used for scientific research and/or transition responsibility to permanent archives
  - Research archive responsibilities persist well beyond lives of missions
  - NASA works with USGS and NARA for long-term preservation
  - NASA has to ensure data and other critical items are preserved and made available to permanent archival agencies
- General requirements
  - No loss of bits
  - Discoverability and accessibility
  - Readability
  - Understandability
  - Usability
  - Reproducibility of results
- NASA has developed Preservation Content Specifications for Earth Science Data
- NASA is participating in Earth Science Information Partners (ESIP) Data Stewardship Committee, on an “emerging” Provenance and Context Content Standard

Categories of Content to be Preserved



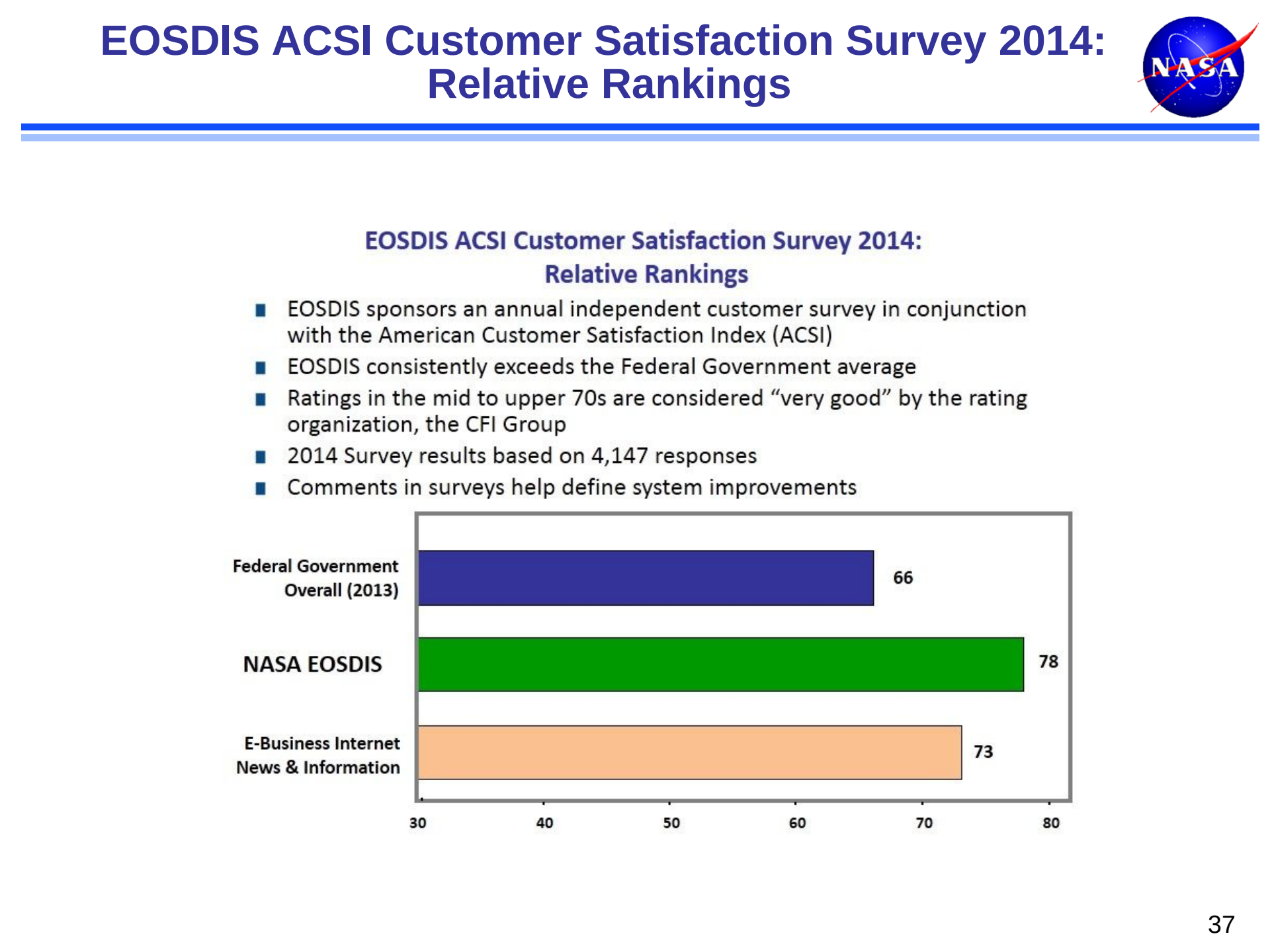
1. **Preflight/Pre-Operations:** Instrument/Sensor characteristics including pre-flight/pre-operations performance measurements; calibration method; radiometric and spectral response; noise characteristics; detector offsets
2. **Science Data Products:** Raw instrument data, Level 0 through Level 4 data products and associated metadata
3. **Science Data Product Documentation:** Structure and format with definitions of all parameters and metadata fields; algorithm theoretical basis; processing history and product version history; quality assessment information
4. **Mission Data Calibration:** Instrument/sensor calibration method (in operation) and data; calibration software used to generate lookup tables; instrument and platform events and maneuvers
5. **Science Data Product Software:** Product generation software and software documentation
6. **Science Data Product Algorithm Input:** Any ancillary data or other data sets used in generation or calibration of the data or derived product; ancillary data description and documentation
7. **Science Data Product Validation:** Records, publications and data sets
8. **Science Data Software Tools:** product access (reader) tools.

Conclusion

- **Success of EOSDIS has been based on its ability to meet and adapt to needs of diverse Earth science communities**
  - **>20 years of diverse science data centers' operation to meet the needs of a growing user community**
  - **>16 years of support for EOS missions (starting with TRMM)**
  - **Working with new missions (e.g., EVS-1 aircraft investigations, EV-2 missions, SMAP, ICESat-2, OCO-2)**
- **Some key areas in which improvements are required and incremental progress is being made:**
  - **Ease of discovery and access**
  - **Cross-organizational interoperability**
  - **Data inter-use**
  - **Ease of collaboration**
  - **Ease of citation of datasets**
  - **Preservation of provenance and context and making them conveniently available to users**

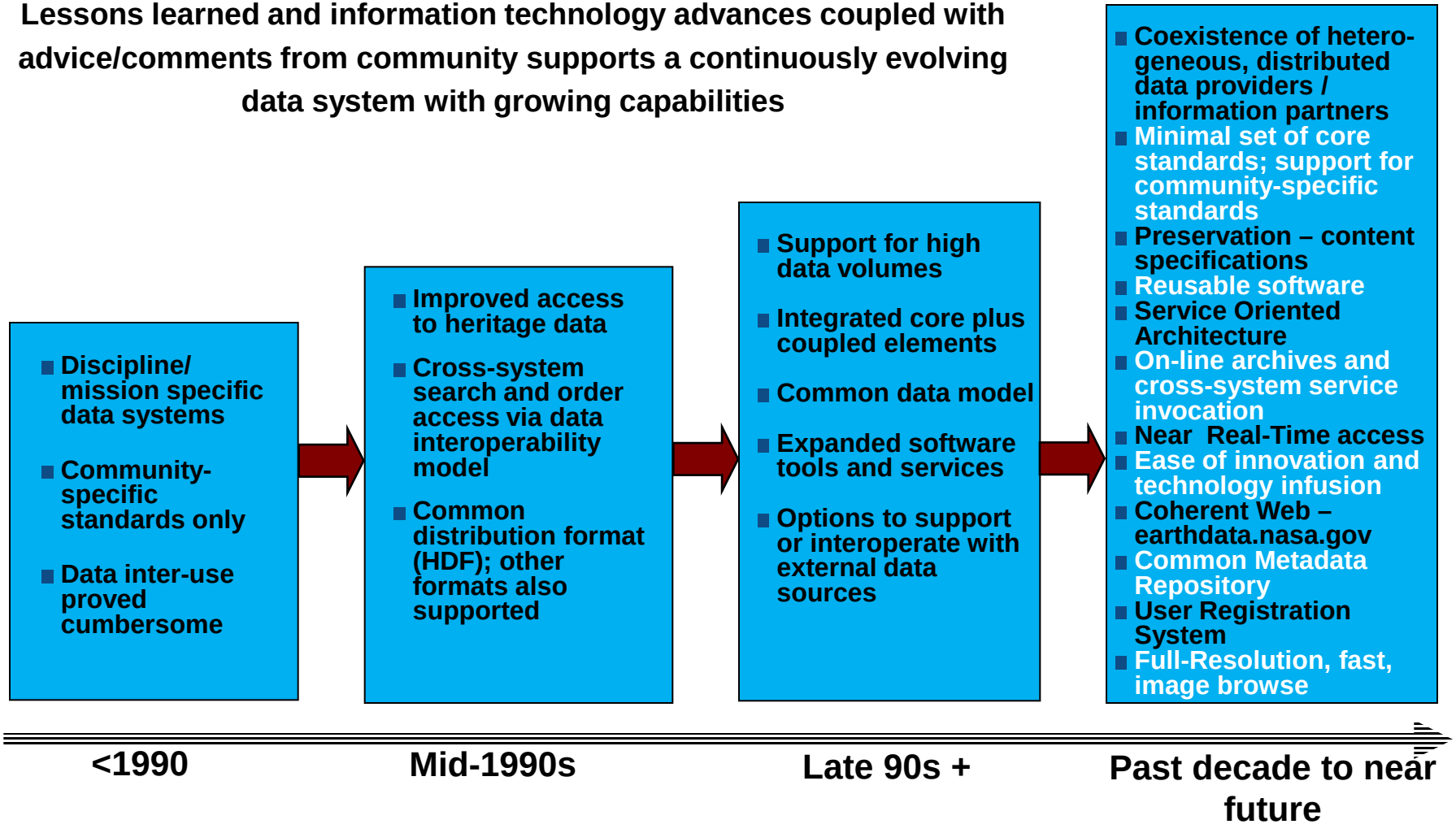
BACK-UP

EOSDIS ACSI Customer Satisfaction Survey 2014: Relative Rankings



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



EOSDIS Acronyms

ACSI	American Customer Satisfaction Index	GES	Goddard Earth Sciences	NRL	Naval Research Laboratory
ACRIM	Active Cavity Radiometer Irradiance Monitor	GHRC	Global Hydrology Resource Center	NSIDC	National Snow and Ice Data Center
AIRS	Atmospheric Infrared Sounder	GLAS	Geoscience Laser Altimeter System	OBPG	Ocean Biology Processing Group
AMSR-E	Advanced Microwave Scanning for EOS	GMAO	Global Modeling and Assimilation Office	OGC	Open Geospatial Consortium
API	Application programming interface	GMU	George Mason University	OMI	Ozone Monitoring Instrument
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer	GRACE	Gravity Recovery and Climate Experiment	ORNL	Oak Ridge National Laboratory
ASDC	Atmospheric Sciences Data Center	GSFC	Goddard Space Flight Center	PB	Peta Byte
ASF	Alaska Satellite Facility	HDF	Hierarchical Data Format	PO.DAAC	Physical Oceanography DAAC
AMSU	Advanced Microwave Sounding Unit	HIRDLS	High Resolution Dynamics Limb Sounder	RBD	Rate Buffered Data
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations	HSB	Humidity Sounder for Brazil	SAGE	Stratospheric Aerosol and Gas Experiment
CDDIS	Crustal Dynamics Data Information System	IWGDD	Interagency Working Group on Digital Data	SAR	Side Aperture Radar
CERES	Clouds and the Earth's Radiant Energy System	JPL	Jet Propulsion Laboratory	SEDAC	Socioeconomic Data and Applications Center
CEOS	Committee on Earth Observation Satellites	LAADS	Level 1 and Atmosphere Archive and Distribution System	SIM	Spectral Irradiance Monitor
CFI	Claes Fornell International	LANCE	Land, Atmosphere Near-real-time Capability for EOS	SIPS	Science Investigator-led Processing Systems
DAAC	Distribute Active Archive Center	LIS	Lightning Imaging Sensor	SNOW-I	Search 'N Order Web Interface
DISC	Data and Information Services Center	LP DAAC	Land Processes DAAC	SOLSTICE	Solar Stellar Comparison Experiment
ECHO	EOS ClearingHOuse	ManLan	Manhattan Landing (high performance exchange point in New York City)	SORCE	Solar Radiation and Climate Experiment
ECS	EOSDIS Core System	MISR	Multi-angle Imaging SpectroRadiometer	SPoRT	Short-term Prediction Research and Transition Center
EDOS	EOS Data and Operations System	MLS	Microwave Limb Sounder	TB	Tera Byte
EOC	EOS Operations Center	MODAPS	MODIS Data Processing System	TES	Tropospheric Emission Spectrometer
EOS	Earth Observing System	MODIS	Moderate Resolution Imaging Spectroradiometer	TIM	TRMM Microwave Imager
EOSDIS	EOS Data and Information System	MOPITT	Measurements of Pollution in the Troposphere	TRMM	Tropical Rainfall Measuring Mission
ESDIS	Earth Science Data and Information System	MSFC	Marshall Space Flight Center	UMBC	University of Maryland, Baltimore County
ESIP	Federation of Earth Science Information Partners	NASA	National Aeronautics and Space Administration	USGS	U.S. Geological Survey
ESSI	Earth and Space Science Informatics	NGIX	Next Generation Internet Exchange	WGISS	Working Group on Information Systems and Services
FGDC	Federal Geographic Data Committee	NISN	NASA Integrated Services Network	WIST	Warehouse Inventory Search Tool
FRGP	Front Range GigaPOP	NITRD	Networking and Information Technology Research and Development	XPS	XUV Photometer System
gbps	Giga bits per second	NPP	NPOESS Preparatory Project		
GCMD	Global Change Master Directory	NPOESS	National Polar-orbiting Operational Environmental Satellite System		