

# **GHRC User Working Group Meeting WELCOME**

...

September 25-26, 2014  
Huntsville, AL



# DEATH BY POWERPOINT

Slow and painful.

fakeposters.com

Source: <http://www.fakeposters.com/posters/death-powerpoint/>

# Agenda – Day 1

0800 - 0830	<i>Coffee; Meet and Greet</i>	
0830 - 0845	GHRC Welcome UAH welcome	Ramachandran Graves
0845 - 0945	Opening Comments <ul style="list-style-type: none"> <li>• Meeting Overview &amp; Logistics</li> <li>• View from HQ</li> <li>• View from ESDIS</li> </ul>	Harrison Kakar Berrick
0945 - 1030	GHRC DAAC Overview Presentation <ul style="list-style-type: none"> <li>• Management &amp; Organization</li> <li>• Vision</li> </ul> DISCUSSION	Conover Ramachandran
1030 - 1100	<i>Break - SPoRT Tour</i>	Zavodsky
1100 - 1200	Infrastructure <ul style="list-style-type: none"> <li>• Hardware and Processing Automation</li> <li>• Data Search, Access and Order Tools</li> </ul> DISCUSSION <ul style="list-style-type: none"> <li>• Web presence</li> </ul> DISCUSSION	Ellett, Beaumont Harrison  Smith, Flynn
1200 - 1300	<i>Working Lunch</i>	
1300 - 1430	Data and Related Tools <ul style="list-style-type: none"> <li>• Satellite and Airborne Microwave Sensors <ul style="list-style-type: none"> <li>○ Data overview</li> <li>○ Regional Air Sea Interactions Tool</li> </ul> </li> </ul> DISCUSSION <ul style="list-style-type: none"> <li>• Field Campaigns <ul style="list-style-type: none"> <li>○ Data overview <ul style="list-style-type: none"> <li>▪ Hurricane Science</li> <li>▪ GPM Ground Validation</li> </ul> </li> <li>○ Mission Coordination Portal</li> <li>○ HS3 Data System</li> </ul> </li> </ul> DISCUSSION	Drewry Keiser  Bugbee Hawkins He Maskey
1430 - 1500	<i>Break</i>	

# Agenda - Day 1

## (cont'd)

1500 - 1700	<p>UWG Member presentations (10 minutes each)</p> <p>Microwave Research</p> <ul style="list-style-type: none"> <li>Assessing Global Water and Energy Budgets</li> <li>Merging retrievals for passive microwave imagers and sounders</li> </ul> <p>Lightning Research</p> <ul style="list-style-type: none"> <li>Lightning research at MSFC</li> <li>Monitoring the performance of space-based optical lightning sensors</li> <li>Multi-sensor data workflows for lightning science</li> </ul> <p>Hurricane Science</p> <ul style="list-style-type: none"> <li>Global distributions of various types of precipitation systems from radar and passive microwave observations</li> <li>The Cyclone Global Navigation System (CYGNSS) Earth Venture Mission</li> <li>Contributions of GHRC datasets to the Zipser Tropical Convection Research Group</li> </ul> <p>GPM Ground Validation</p> <ul style="list-style-type: none"> <li>GPM Field Campaigns</li> </ul> <p>Applications</p> <ul style="list-style-type: none"> <li>Integration of Earth Science Research and Education at UAH</li> <li>SPoRT/MSFC Applied Science: Disaster Activities</li> </ul>	<p>Kummerow Duncan</p> <p>Blakeslee Buechler</p> <p>Bruning</p> <p>Liu</p> <p>Ruf</p> <p>Zawislak</p> <p>Wolff</p> <p>Griffin</p> <p>Molthan</p>
1700-1730	Summary, Actions and Wrap-up	Ramachandran
1730	<i>Adjourn</i>	
1830	<i>Group Dinner - 1892 East</i>	

# Agenda – Day 2

0800 - 0830	<i>Coffee; Meet and Greet</i>	
0830 - 0845	Welcome, recap, plan for day	Ramachandran
0845 - 0930	Lightning Data and Tools <ul style="list-style-type: none"> <li>• Data overview</li> <li>• LIS Space Time Search</li> <li>• GLM Validation Tool</li> </ul> DISCUSSION	Regner
0930 -1000	AMSR SIPS highlights <ul style="list-style-type: none"> <li>• LANCE AMSR2</li> <li>• AMSR-E Provenance Browser</li> </ul> DISCUSSION	Lin McEniry
1000 - 1020	ESDIS initiatives DISCUSSION	Conover
1020 - 1030	Charge to UWG	Ramachandran
1030 - 1100	<i>Break - SERVIR Tour</i>	Irwin
1100 - 1330	Executive Session - discussion and report <i>(working lunch provided)</i>	UWG
1330 – 1430	Board Advice to DAAC - prioritization of current work and future efforts	All
1430 – 1500	Summary of Actions and Meeting Wrap-up	Ramachandran
1500	<i>Adjourn</i>	

# GLOBAL HYDROLOGY RESOURCE CENTER

A NASA Distributed Active Archive Center

**Rahul Ramachandran**

DAAC Manager

[rahul.ramachandran@nasa.gov](mailto:rahul.ramachandran@nasa.gov)

**Helen Conover**

GHRC Operations Manager

[hconover@itsc.uah.edu](mailto:hconover@itsc.uah.edu)

Presented at the GHRC User Working Group Meeting  
September 25-26, 2014





# Global Hydrology Resource Center

- **Full service data center** providing data ingest, routine and custom processing, archive, distribution, user support, and science data services
- **Collaboration** between NASA and the University of Alabama in Huntsville to **infuse advanced information technologies** to a variety of **science data projects**
- Global **lightning data from space**, airborne and ground based observations from **hurricane science field campaigns** and Global Precipitation Mission (**GPM**) **ground validation** experiments, and **satellite passive and active microwave** products

<http://ghrc.nsstc.nasa.gov/>



# What we do

## Data Stewardship



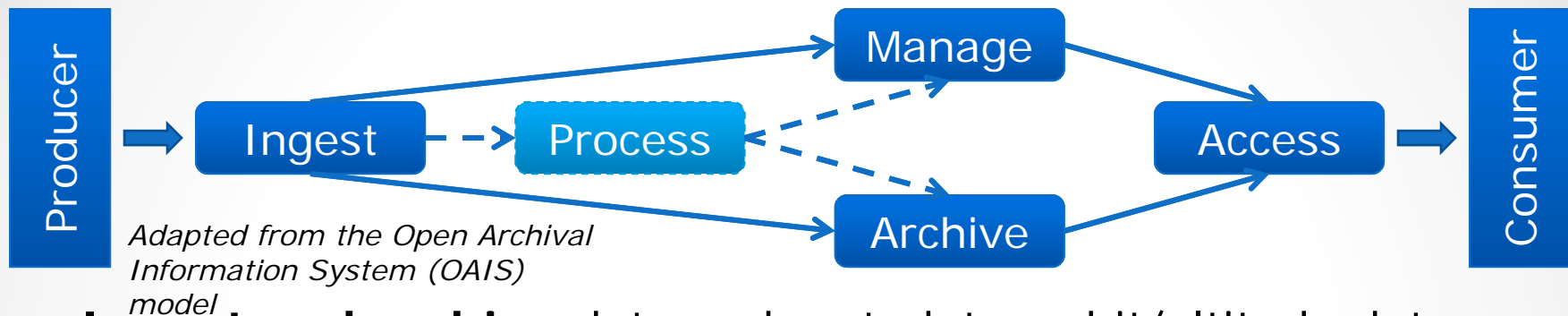
NASA's Earth science data stewards for scientific, educational, commercial and governmental communities, with a focus on data for the global hydrologic cycle



# History

- **Marshall DAAC** was established in 1991 at the beginning of NASA's EOSDIS program
  - Based on the WetNet project led by Michael Goodman and a local science data management effort led by Sara Graves.
  - Science focus was passive microwave and lightning data
- **LIS Enhanced Science Computing Facility (E-SCF)** was established 1997 to manage data from the Lightning Imaging Sensor on TRMM
  - Co-branded as **Global Hydrology Resource Center**
  - Funding through the MSFC lightning science team
  - Supplemental funding through other science projects (e.g., the Hurricane Science Program for specific field campaigns)
- **GHRC DAAC** was added to the NASA Earth Science Data and Information Systems (ESDIS) project for core funding in 2009
- **AMSR-E SIPS at GHRC** was established in 1998 to generate standard products from the AMSR-E instrument on Aqua. Near-real time processing for **LANCE** was added in 2010.

# Data Center Operations



- **Ingest and archive** data and metadata, orbit/altitude data, documents, algorithms, instrument and spacecraft history, ancillary data from external sources for production.
- **Processing** including science product generation and reformatting, algorithm integration and test, interfaces with external providers (EOSDIS, other data centers).
- **Data discovery and access services** include direct online access to most data products, an online search and order system, registration of all data in NASA Earth science data catalogs, and support for a variety of data access web services
- **Data distribution and user services** including processing orders (subscriptions and on-demand), tracking orders across system, prioritizing based on resource management policy.

# GHRC Staffing Profile

*Except for DAAC Manager, all GHRC staff are matrixed from UAH's Information Technology and Systems Center.*

## **Mission and Science Support**

- Metadata Development
- Documentation
- Science Team Interactions

## **User Services**

- Customer Interactions
- Web Site and Social Media

## **Operations**

- Ingest, Processing and Archive Management
- Systems Testing

## **Engineering**

- Systems Engineering
- Software Engineering and Development
- HS3 Data System Planning
- Evolution of Existing Systems

## **Infrastructure**

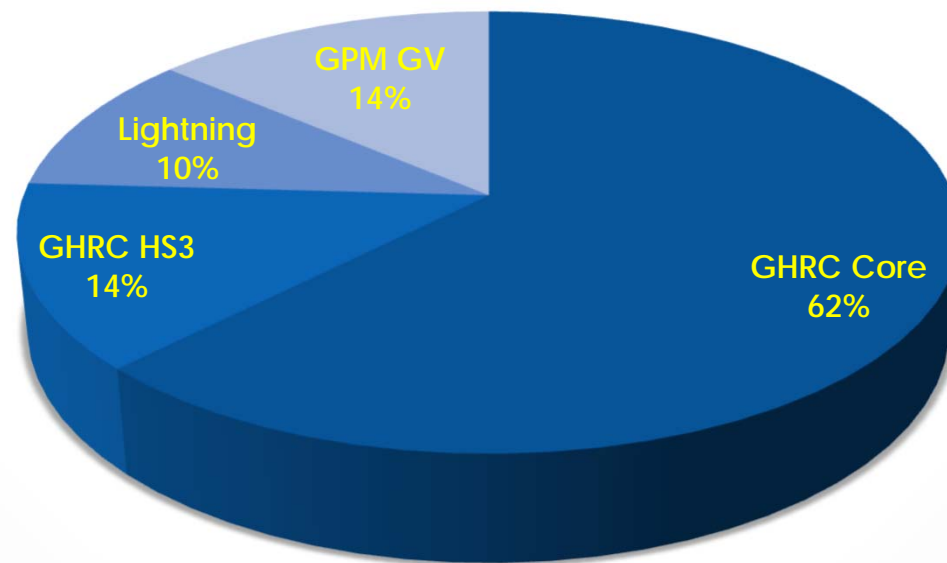
- Systems Administration
- Database Administration

## **IT Security**

## **Project Management**

# Shared Resources

- GHRC leverages supplemental funding from science projects to provide data management services using GHRC infrastructure
- GHRC also shares staff with the AMSR SIPS

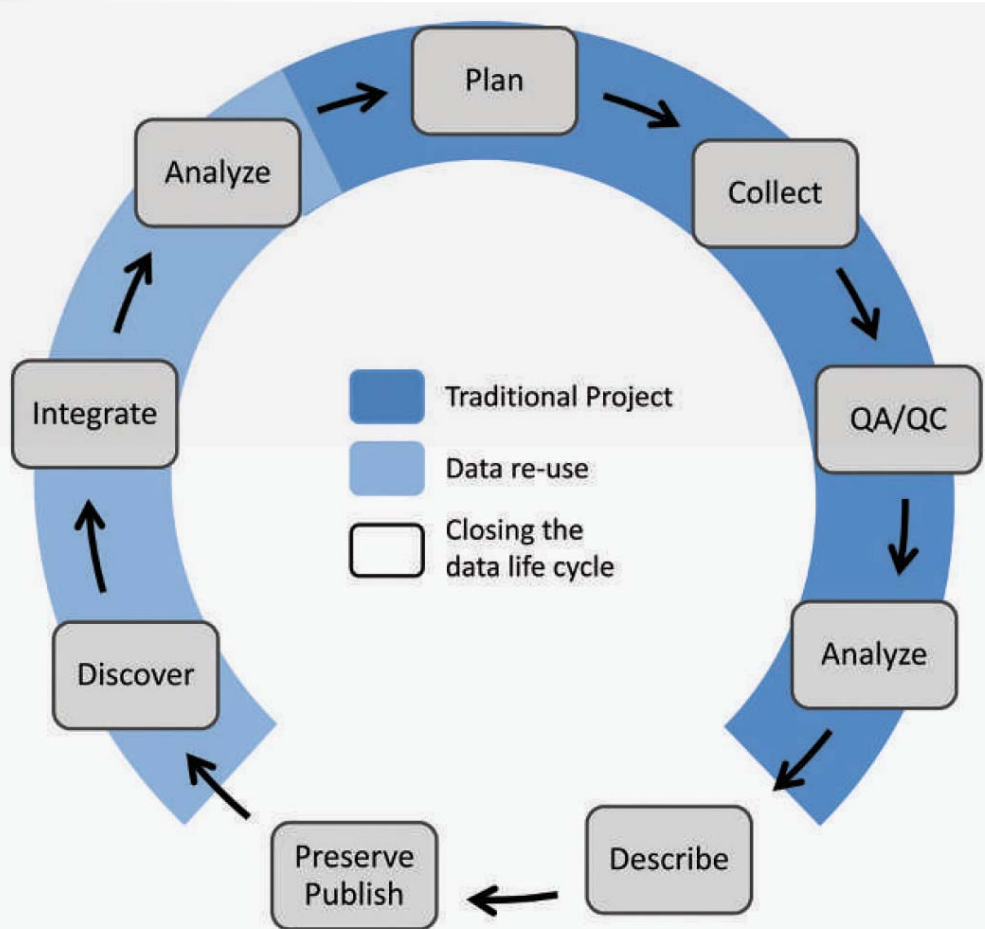


*Combined funding snapshot for 2014*

# Importance of (Open) Data

- Fair Access to Science and Technology Research Act (FASTR) introduced in both the Senate and the House in 2013.
- OSTP memorandum :  
“directs each Federal agency with over \$100 million in annual conduct of research and development expenditures to develop a plan to support increased public access to the results of research funded by the Federal Government. This includes any results published in peer-reviewed scholarly publications that are based on research that directly arises from Federal funds . . .” (OSTP 2013, 2).
- Explicitly states that “such results include peer-reviewed publications and digital data.”

# Role of DAAC's in the Data Life Cycle



- Provides two steps needed to complete the data lifecycle
- Enables data to retain value past the life of the project and creates new research/application opportunities



# GHRC Mission Statement

- To serve as NASA's Earth science *data stewards* for scientific, educational, commercial and governmental communities, with a focus on data for the global hydrologic cycle
  - Hydrologic Cycle
  - Severe Weather Interactions
  - Lightning
  - Atmospheric Convection
- To provide *knowledge augmentation services* encompassing tools, infrastructure, user support, and expertise to our stakeholders

# What we do

## Knowledge Augmentation Services

**PROVENANCE**  
Make the preserved data/information available to all our stakeholder communities with traceability to support authenticity  
(*AMSR-E Provenance*)

**DATA DISCOVERY**  
Develop new tools for data discovery, curation and aggregation  
(*LIS Interactive Browse*)



**FIELD CAMPAIGN INFRASTRUCTURE**  
Create specialized portals for managing field campaigns and collecting data  
(*Field Campaign Portal*)

**DATA USE**  
Develop new tools for access, analysis and visualization  
(*HS3 Data System, GLM Validation Tool, RAS*)

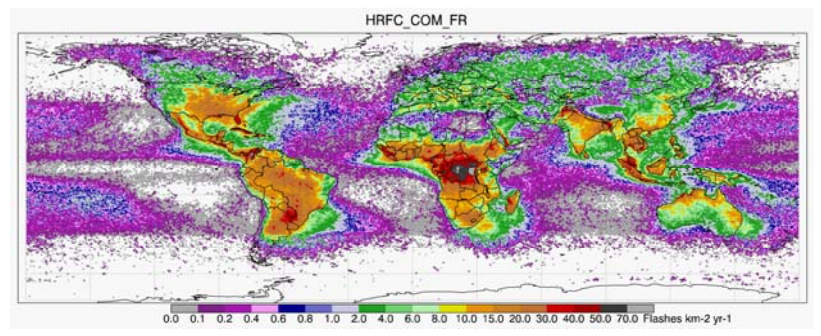
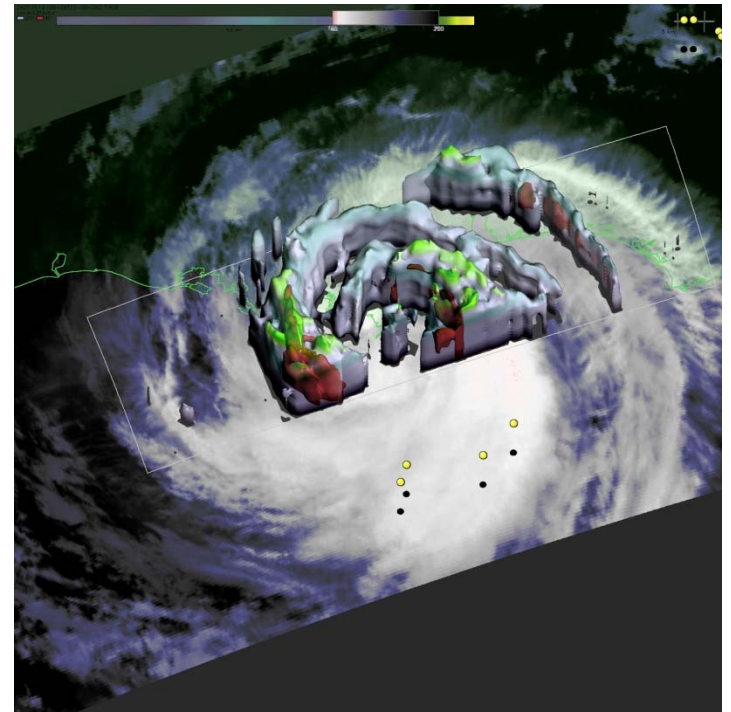
**INFUSING CUTTING EDGE INFORMATICS**  
Research new approaches and technologies and infuse them into operational processes

GHRC provides knowledge augmentation services encompassing tools, infrastructure, user support, and expertise to our stakeholders

# What we serve

## Lightning Data

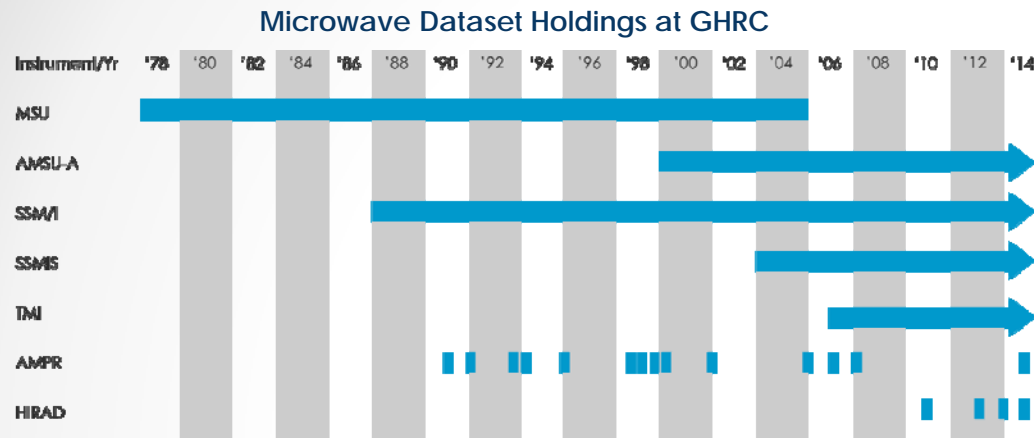
- Responsible for data from the TRMM Lightning Imaging Sensor, plus ancillary lightning data sets utilized by the LIS SCF scientists, since January 1998. A second LIS instrument will fly on the SpaceX rocket to the International Space Station in February 2016.
- Ancillary data –
  - National Lightning Detection Network, electric field mill data from the Kennedy Space Center, global infrared data and ground based radar data
- Precursor satellite instruments -
  - Optical Transient Detector in operation on Microlab-1 from 1995 to 2000
  - Operational Linescan Sensor on Defense Meteorological satellites from 1973 to 1995



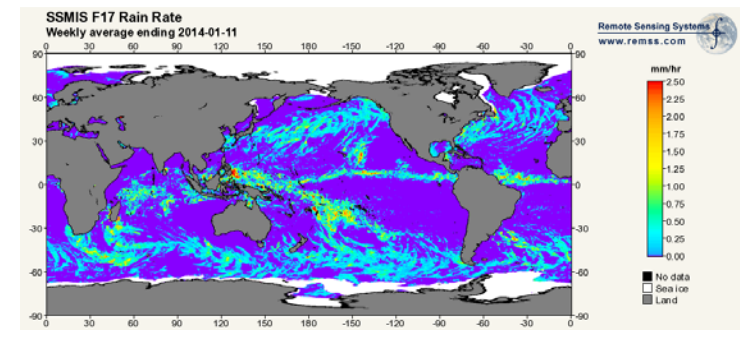
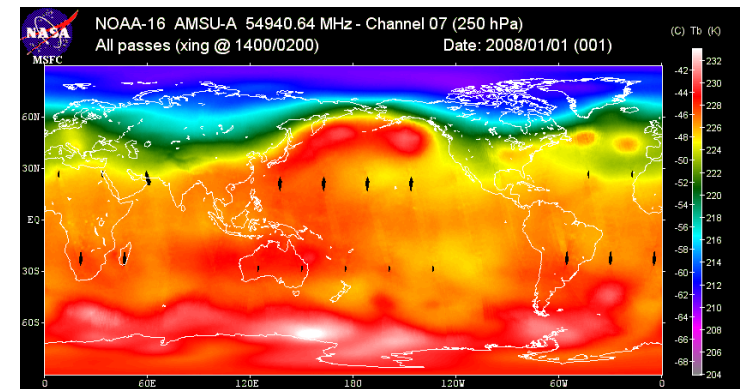
**GHRC is recognized as the National Lightning Archive**

# What we serve

## Microwave Data



- GHRC and its predecessor programs have been ingesting, processing, archiving and distributing microwave data for over 35 years
  - MSU, SSMI, AMSU, AMPR, TMI, AMSR-E
- This climate sensitive data record extends back to 1978 providing an unbroken inventory of climate information that continues today



GHRC is also recognized as one of the primary data centers for microwave data

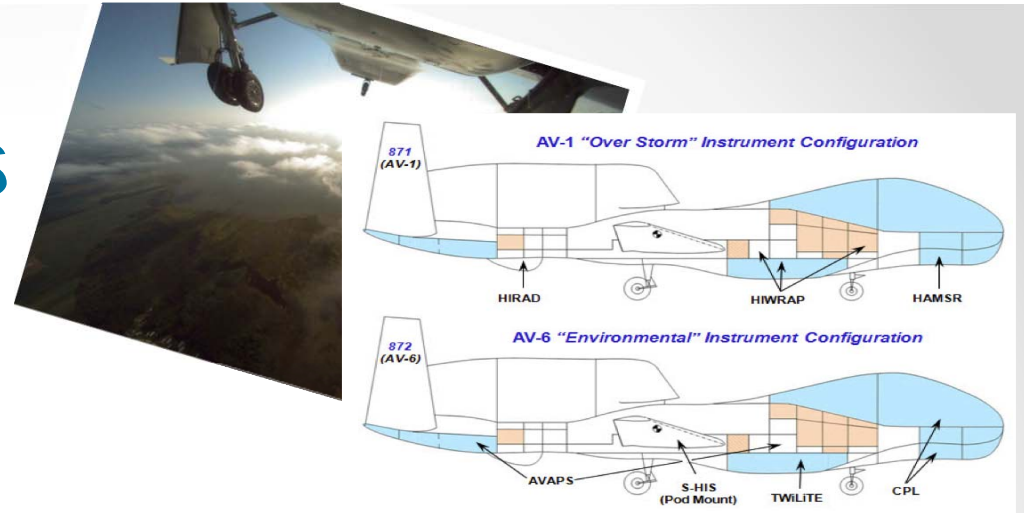


# What we serve

## Field Campaigns

### Hurricane Science

Data from successive field campaigns since 1990 are tied together through common procedures, consistent metadata, and discovery and archival systems making it easy to access data from instruments that have been employed across several missions



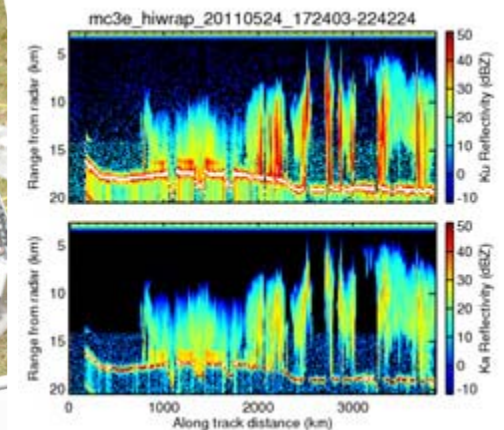
### Hurricane and Severe Storm Sentinel (HS3)

Five-year mission to investigate the processes that underlie hurricane intensity change in the Atlantic Ocean basin and will utilize two Global Hawks

GHRC is recognized as one of the main data centers for Hurricane Science data

### Global Precipitation Measurement Mission (GPM) Ground Validation (GV)

Ground and airborne precipitation datasets supporting physical validation of satellite-based precipitation retrieval algorithms



GHRC is set up to manage a large number of episodic, heterogeneous datasets and can handle the "long tail" of science data

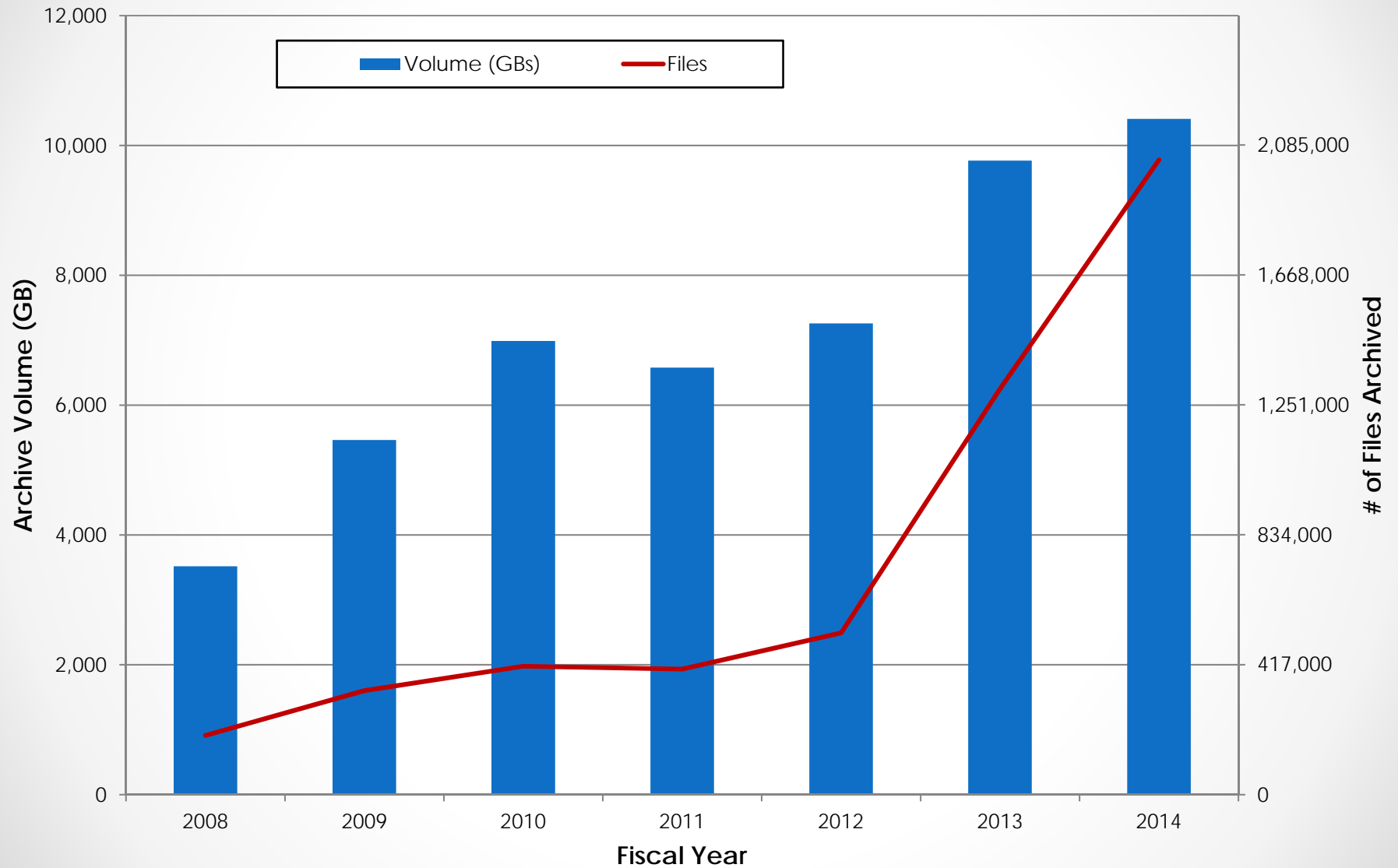
# GHRC By the Numbers



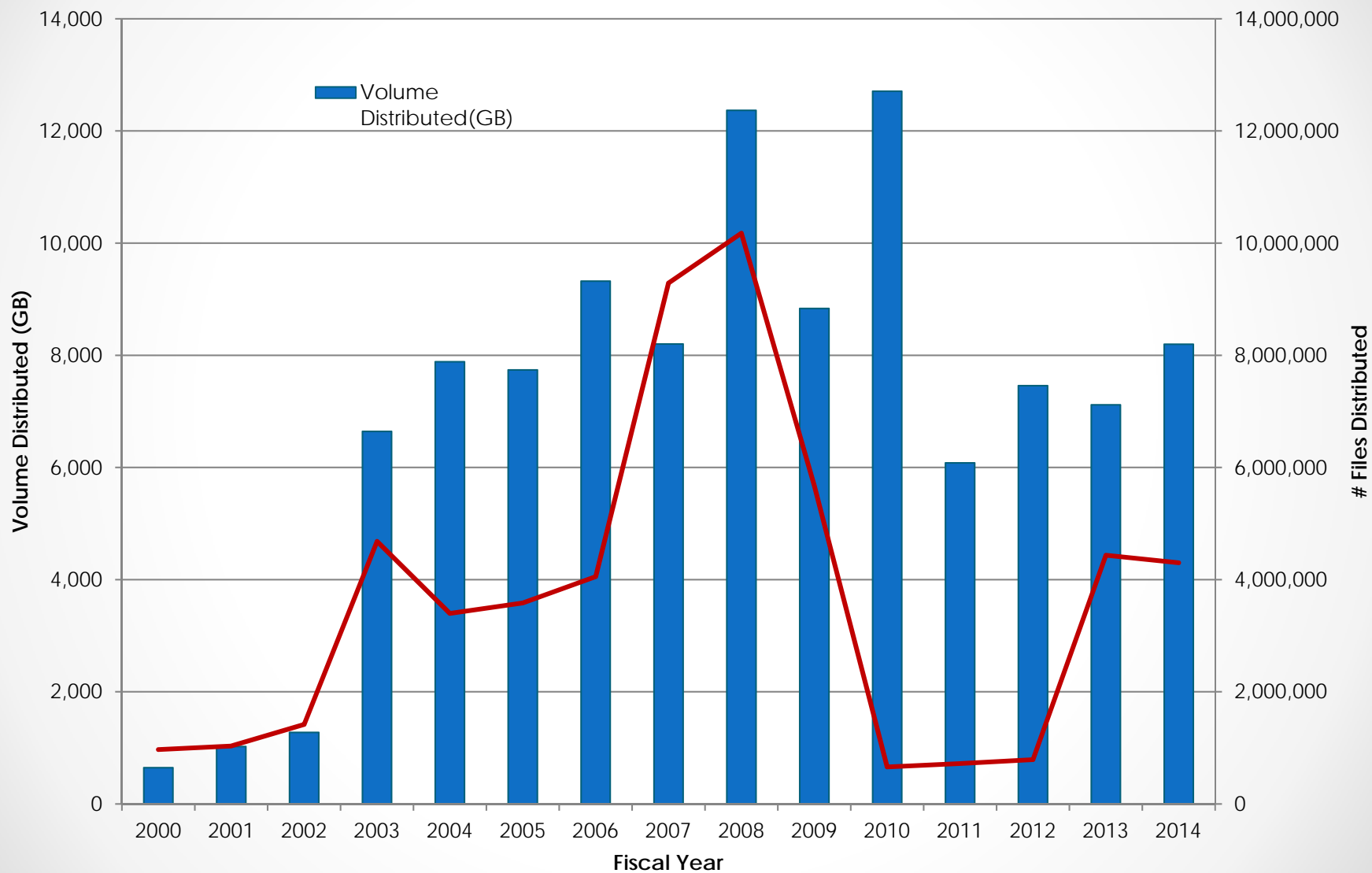
- Registered Users – 1341
- Data sets
  - 291 Public
  - ~ 8% of total ESDIS holdings (3666)
  - 8 Limited visibility
  - 34 input data streams only used in processing to produce the final products
- Granules
  - ~ 2 Million (Archived since 94)
- Archive size - ~ 10 TB
  - HS3 will add 60 TB
- Distribution
  - ~82 million files since 94



## GHRC Yearly Cumulative Archive (Oct 1, 2008 - Aug 29, 2014)



## GHRC Yearly Data Distribution



# Data Impact beyond Science Teams

Tellus

SERIES A  
DYNAMIC  
METEOROLOGY  
AND OCEANOGRAPHY

PUBLISHED BY THE INTERNATIONAL METEOROLOGICAL INSTITUTE IN STOCKHOLM

*Tellus* (2010), 62A, 393–410

Printed in Singapore. All rights reserved

© 2010 The Authors

Journal compilation © 2010 Blackwell Munksgaard

TELLUS

*Science in China Ser. D Earth Sciences* 2005 Vol.48 No.2 219–229

219

## Improving passive microwave sea ice concentration algorithms for coastal areas: applications to the Baltic Sea

By NINA MAAß\* and LARS KALESCHKE, *Institute of Oceanography, University of Hamburg, 20146 Hamburg, Germany*

**Title:** Preliminary Lightning Observations over Greece

**Author:** Themis G. Chronis

**Address:** Hellenic Center for Marine Research (HCMR), Institute of Marine Biology, 16.7 km Athens - Sounion Av., P. O 19013 Anavyssos

**Address:** 16.7 km Athens - Sounion Av., P. O 19013 Anavyssos

## Climatological distribution of lightning density observed by satellites in China and its circumjacent regions

MA Ming, TAO Shanchang, ZHU Baoyou & LÜ Weitao

*Science in China Series D: Earth Sciences*  
February 2005, Volume 48, Issue 2, pp 219–229

## Climatological distribution of lightning density observed by satellites in China and its circumjacent regions

Ming Ma, Shanchang Tao, Baoyou Zhu, Weitao Lü



## Atmospheric Research

Volume 91, Issues 2–4, February 2009, Pages 438–452

13th International Conference on Atmospheric Electricity — ICAE 2007

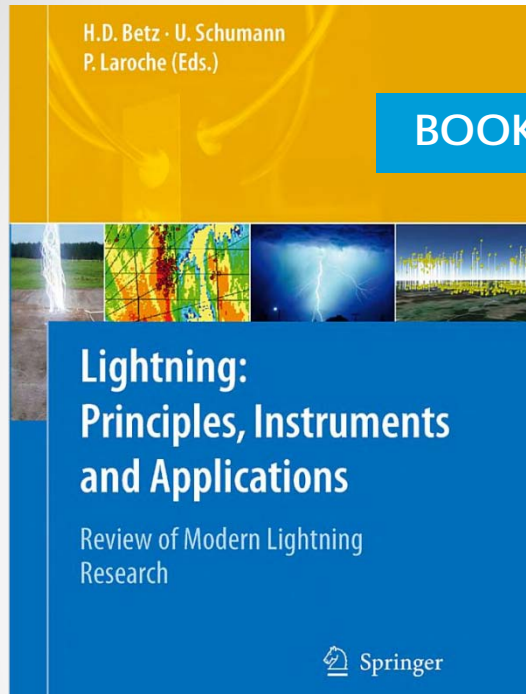


## A comparison of lightning activity and convective indices over some monsoon-prone areas of China

Jianhua Dai<sup>a, b</sup>, Yuan Wang<sup>a</sup>, Lei Chen<sup>b</sup>, Lan Tao<sup>b</sup>, Jianfeng Gu<sup>b</sup>, Jianchu Wang<sup>c</sup>, Xiaodong Xu<sup>c</sup>, Hong Lin<sup>b</sup>, Yudan Gu<sup>c</sup>

GHRC data is used to address specific regional needs all over the world

# Data Impact beyond Science Teams



BOOKS



## Terrestrial Gamma-Ray Flashes Observed up to 20 MeV

David M. Smith *et al.*  
*Science* **307**, 1085 (2005);  
DOI: 10.1126/science.11074

New science areas



## Direct Measurements of the Convective Recycling of the Upper Troposphere

Timothy H. Bertram *et al.*  
*Science* **315**, 816 (2007);  
DOI: 10.1126/science.1134548

WILDERNESS & ENVIRONMENTAL MEDICINE, **23**, 260–269 (2012)

WILDERNESS MEDICAL SOCIETY PRACTICE GUIDELINES

## Wilderness Medical Society Practice Guidelines for the Prevention and Treatment of Lightning Injuries

VOLUME 11

JOURNAL OF HYDROMETEOROLOGY

AUGUST 2010 Eric Johnson, MD; Scott E. McIntosh

Science Applications

## The Hydrology and Hydrometeorology of Flooding in the Delaware River Basin

JAMES A. SMITH, MARY LYNN BAECK, AND GABRIELE VILLARINI  
*Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey*

WITOLD F. KRAJEWSKI  
*IIHR-Hydropscience and Engineering, The University of Iowa, Iowa City, Iowa*

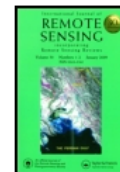
(Manuscript received 13 November 2009, in final form 22 March 2010)

International Journal of Remote Sensing

Volume 31, Issue 13, 2010

Special Issue: Satellite Observations of the Wenchuan Earthquake of 12 May 2008

Select Language ▼  
Translator disclaimer



Precursory signals using satellite and ground data associated with the Wenchuan Earthquake of 12 May 2008

Ancillary data from GHRC is used in many, many papers

# Our Vision for GHRC

- Efficient
  - Minimize any operational redundancies via automation
- Innovative
  - Design, develop and adopt new technologies to minimize cost and maximize productivity of our stakeholders
- Agile
  - Respond to changing needs (science driven/programmatic)
- Active
  - Collaborations with our stakeholders
  - Leadership roles in ES Informatics



<http://inspirationalstorytellers.com/wp-content/uploads/2013/05/future-vision.jpg>



Flat budgets

# Community/Leadership Activities

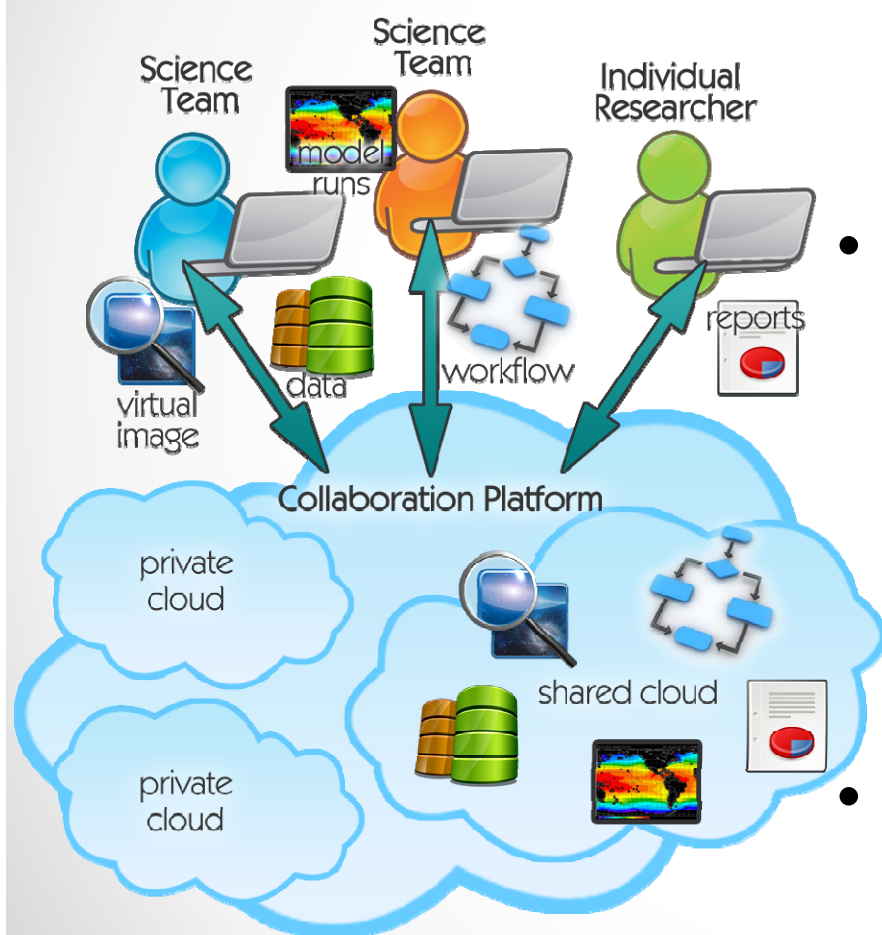
- ESDSWG: focuses on community driven recommendation for Earth Science data system
  - Innovations Lab Working Group (McEniry, Ramachandran)
  - Airborne Working Group (Conover)
  - ASCII for Science Data (Conover)
- ESDIS Standards Office (Conover)
- IEEE GRSS Earth Science Informatics Technical Committee – (Ramachandran)
- AGU Earth and Space Science Informatics (ESSI) Focus Group



# New Projects

- Computational Modeling Algorithms and Cyberinfrastructure (CMAC) program: *Collaborative Workbench to Accelerate Science Algorithm Development* (PI Ramachandran/UAH PI Maskey)
- Advanced Information Systems Technology (AIST) Program: *Automated Event Services: Efficient and Flexible Searching for Earth Sciences Phenomena* (PI Clune GSFC/Co-I Ramachandran)
- HS3 Data and Information System
- White House OSTP led Climate Data Initiative (CDI)
- White House OSTP Big Earth Data Initiative (BEDI)

# Earth Science Collaboration Workbench (CWB)

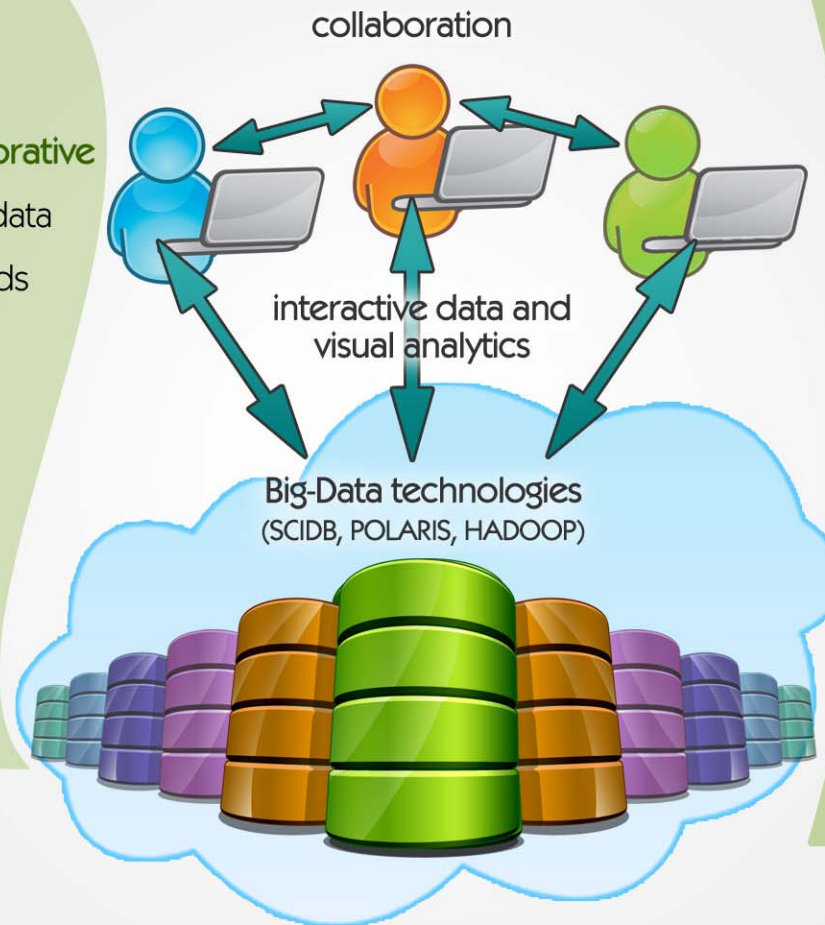


- Augments a scientist's current research environment to allow him or her to easily share diverse data and algorithms
- Leverages technologies such as the cloud and social collaboration frameworks for scalable and controlled collaboration
  - Open source Eclipse framework, compatible with widely used scientific analysis tools such as IDL and Python.
- Misc.:
  - GLM Validation and Verification Tool
  - Provenance Service

# Automated Event Services

Utilize **Big-Data** technologies to...

- Enable **interactive** and **collaborative** scientific **data analysis** on big data
- Share data and analysis methods seamlessly,  
...in order to...
- Relieve scientists from data management,
- Empower scientists to focus on science, and
- Boost science productivity.



- 1 **Identify occurrences** (events) of phenomena
  - Entities in the 4D spatiotemporal space
- 2 **Associate** additional relevant data with events.
- 3 **Characterize phenomena** with defining features extracted from data.
- 4 **Correlate** defining features of various phenomena in both space and time.
- 5 **Improve predictions** of future events using correlations among phenomena for better decision making.

Big-Data  
Vision

Technology  
Infrastructure

Science  
Enablement

# Climate Data Initiative

- NASA is leading the Climate Data Initiative being coordinated by the Council on Environmental Quality (CEQ) and the Office of Science and Technology Policy (OSTP).
  - Identify and make interoperable relevant data from multiple interagency sources to support climate
  - Facilitate the integration and better use of data for decision support and actionable science information
  - Make these data more accessible, discoverable, and usable for purposes other than which they were originally collected
- NASA has formed a Data Coordination team consisting of personnel from GHRC with appropriate expertise to support these goals.

The screenshot shows the DATA.GOV website interface. At the top, there's a navigation bar with 'DATA', 'TOPICS', 'IMPACT', 'APPLICATIONS', 'DEVELOPERS', and 'CONTACT'. Below this, a blue header bar says 'CLIMATE'. A sub-navigation bar lists 'Themes', 'Data', 'Resources', 'Challenges', 'FAQ', and 'Feedback'. The main content area has a paragraph about finding climate change data. Below that, a 'HIGHLIGHTS' section features 'NASA launches Earth Science Challenges'. The bottom section, '143 datasets found', lists several datasets with brief descriptions and file format icons (XML, CSV, PDF, etc.). On the right side of the dataset list, there are blue arrows pointing right, each labeled 'Federal'.

DATA.GOV

DATA TOPICS IMPACT APPLICATIONS DEVELOPERS CONTACT

CLIMATE

Themes Data Resources Challenges FAQ Feedback

Here you can find data related to climate change that can help inform and prepare America's communities, businesses, and citizens. Initially, in this pilot phase, you can find data and resources related to [coastal flooding](#) and [food resilience](#). Over time, you will be able to find additional data and tools relevant to other important climate-related impacts, including risks to human health, ecosystems, and energy infrastructure. Please share your [feedback](#).

HIGHLIGHTS

**NASA launches Earth Science Challenges**

Through the Open NASA Earth Exchange, NASA is challenging innovators to use data from Earth's satellites in new and creative ways, including for applications that help people understand climate change. [Learn more](#).

[More Highlights](#)

143 datasets found

**MyPyramid Food Raw Data**

Department of Agriculture — MyPyramid Food Data provides information on the total calories; calories from solid fats, added sugars, and alcohol (extras); MyPyramid food group and subgroup...

**Commodity Flow Survey**

US Census Bureau, Department of Commerce — The Commodity Flow Survey provides information on commodities shipped, their value, weight, and mode of transportation, as well as the origin and destination of...

**Global Climate Change**

National Aeronautics and Space Administration — NASA's Eyes on the Earth: Global Climate Change and Global Warming. Current news and data streams about global warming and climate change from NASA's Jet Propulsion...

**Internet Weather Source**

National Oceanic and Atmospheric Administration, Department of Commerce — The National Weather Service (NWS) National Telecommunications Gateway provides weather, hydrologic, and climate forecasts and warnings for the United States, its...

**Current Year National Agriculture Imagery Program (NAIP) Imagery**

Farm Service Agency, Department of Agriculture — Provides the ArcGIS Rest Services URLs for the public facing, most current year 1-meter natural color NAIP web services for the lower 48 states, served by USDA-FSA...



# Big Earth Data Initiative (BEDI)

- BEDI seeks to improve the discoverability, accessibility, and usability of data and information derived from Federal civil Earth observations, making these information products easier for everyone to find and use.
- *GHRC Task*
  - Data available online via services based on open standard protocols
    - Focus on Field Campaigns

# Proposal Submitted

- Submitted to Advanced Information Systems Technology (AIST) Program
  - Developing a Numerical Weather Prediction and Data Dissemination Virtual Appliance to Support Disaster Preparedness, Mitigation, and Response (PI Molthan MSFC/Co-I Ramachandran)
  - DEREChOS: Data Environment for Rapid Exploration and Characterization of Organized Systems (PI Clune GSFC/Co-I Ramachandran)
  - GEODE: GEO Data Engine to Enable Big Data Analytics in Exascale-Computing Era (PI Ramachandran)
  - Illuminating the Darkness: Exploiting untapped data and information resources in Earth Science (PI Ramachandran)
- Pursuing other funding opportunities to build new capacity within GHRC



# UWG Charge



<http://www.marketingbrainfodder.com/files/2012/09/Inspire-300x199.jpg>

- Select a scribe
- Elect a co-chair
  - Executive Session Friday Morning
- Provide prioritized recommendations/suggestions for improvements
- Tell us what we are doing right
- Write and submit a report

# UWG Questions

- Data Stewardship

- Are there additional important ancillary data that need to be in the GHRC catalog?
- How can we make GHRC more visible to the science communities?
  - Science conferences/Meetings

- Knowledge Augmentation Services

- Are there other services that GHRC can provide that will make your research process easier?
  - New means of data discovery
- Should GHRC look at changing the access mechanisms?
  - Accessing data files from machine APIs (libs for python, idl)
  - Cloud based stores (AWS S3 for EC2 computation)
- Need for new tools for data exploration and visualization?