Information Technology and Systems Center

IT Research and Development

- Data Mining and Exploitation
- Remote Sensing Data Analysis
- Cybersecurity
- On-demand Processing
- Sensor Networking Analysis

NSSTC Core Facility

Sara J. Graves. Ph D

Director, Information Technology and Systems Center
University Professor, Computer Science Department
University of Alabama in Huntsville
Director, Information Technology Research Center
National Space Science and Technology Center
256-824-6064
sgraves@itsc.uah.edu

http://www.itsc.uah.edu

Information Technology and Systems Center

Providing tools to access, manage, analyze, and visualize data, information and knowledge

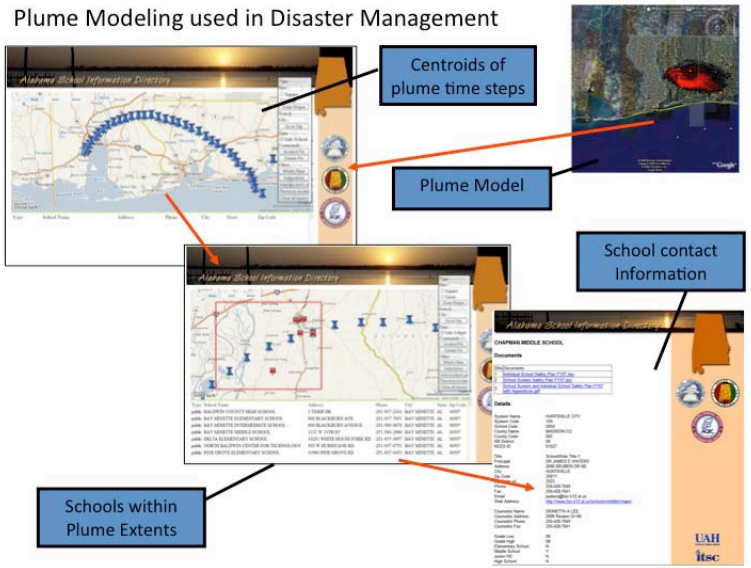
- *Data and information management*
- *Geographic Information Systems*
- *Ontologies, Semantics and Semantic Web*
- *Data mining and decision support*
- *Modeling and simulation*
- *Visualization*
- *Collaborative environments*

Getting the right information to the right people at the right time

- *Coastal science*
- *Climate and weather*
- *National Security*

enhancing the ecological and economic health of the Gulf of Mexico

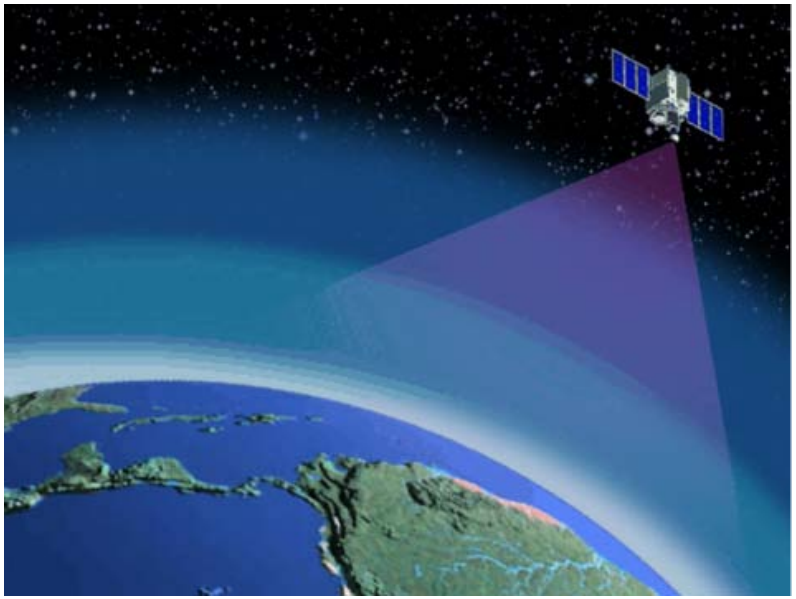
Plume Modeling used in Disaster Management



* Projects funded by multiple agencies

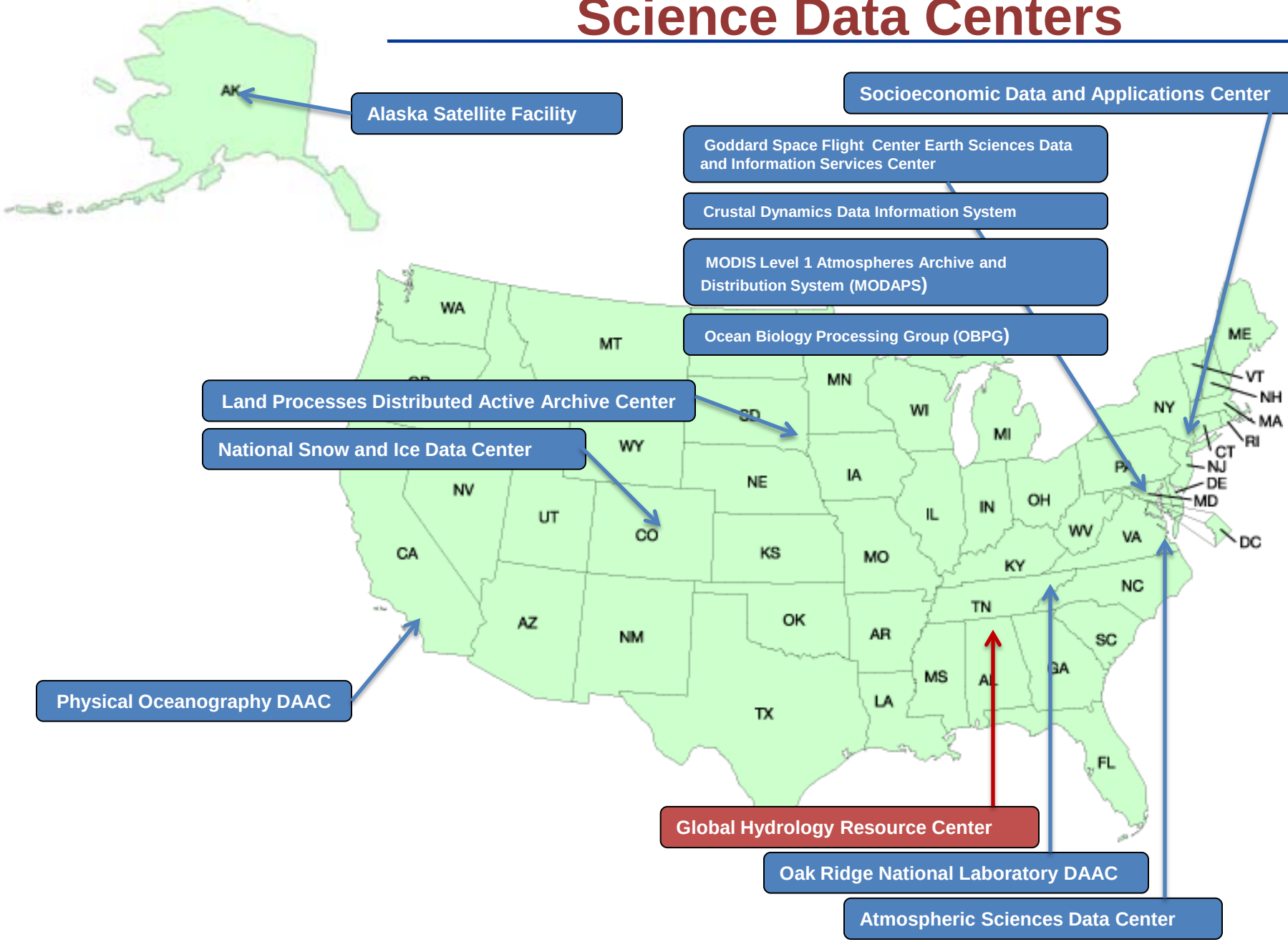
Aspects of Environmental Data Management

- Planning – requirements analysis, standards-based formats, metadata, catalogs, registries
- Provenance
- Ingest and processing
- Integration – with data, products, and tools
- Analysis and Visualization
- Archiving
- User Support
  - Information
  - Data Access and discovery
  - Interoperable Tools and services

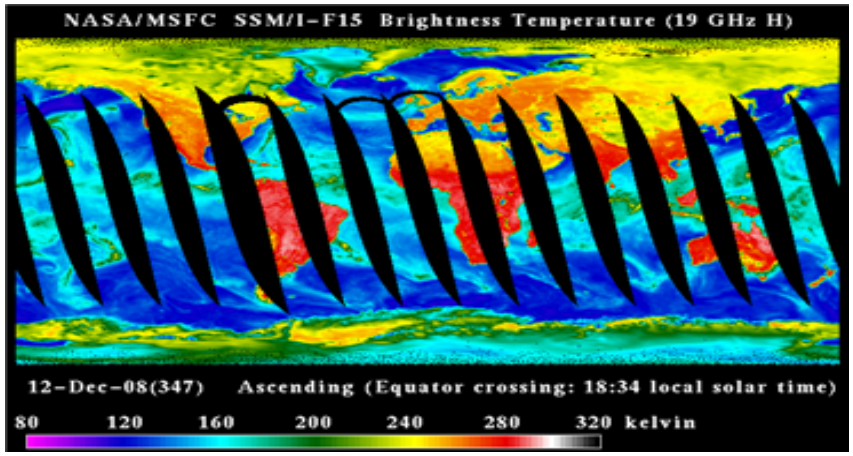


An Overview of the Global Hydrology Resource Center

The GHRC is One of 12 NASA Earth Science Data Centers

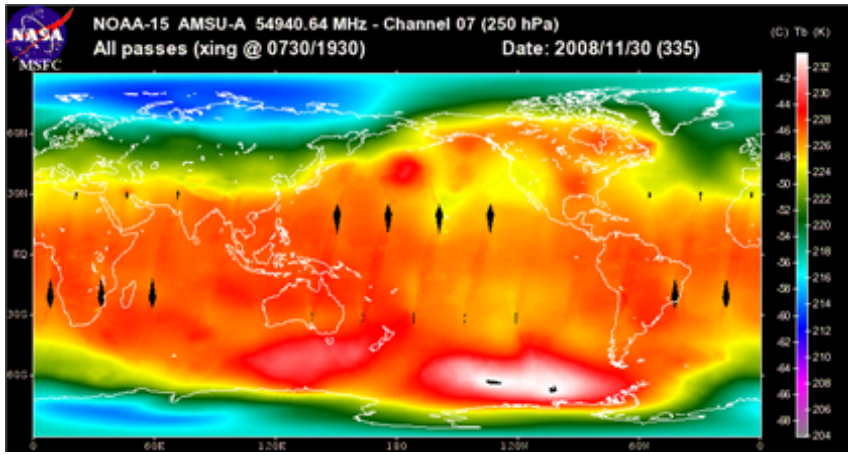


Passive Microwave Data – 15 Year Archive



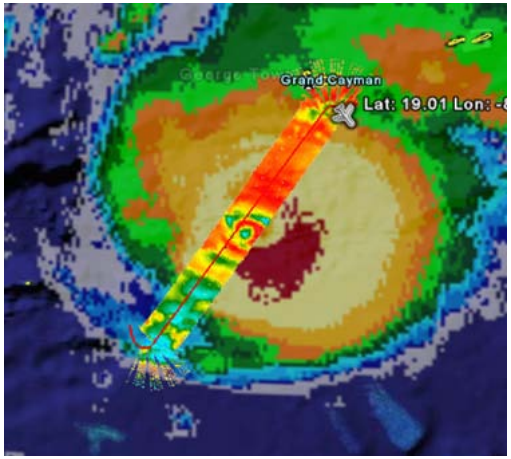
Special Sensor Microwave Imager

Brightness Temperatures
Surface Winds
Water vapor
Cloud Liquid Water & Ice



Advanced Microwave Sounding Unit

Brightness Temperatures
Air Temperature
Temperature Anomalies (Day/Month)

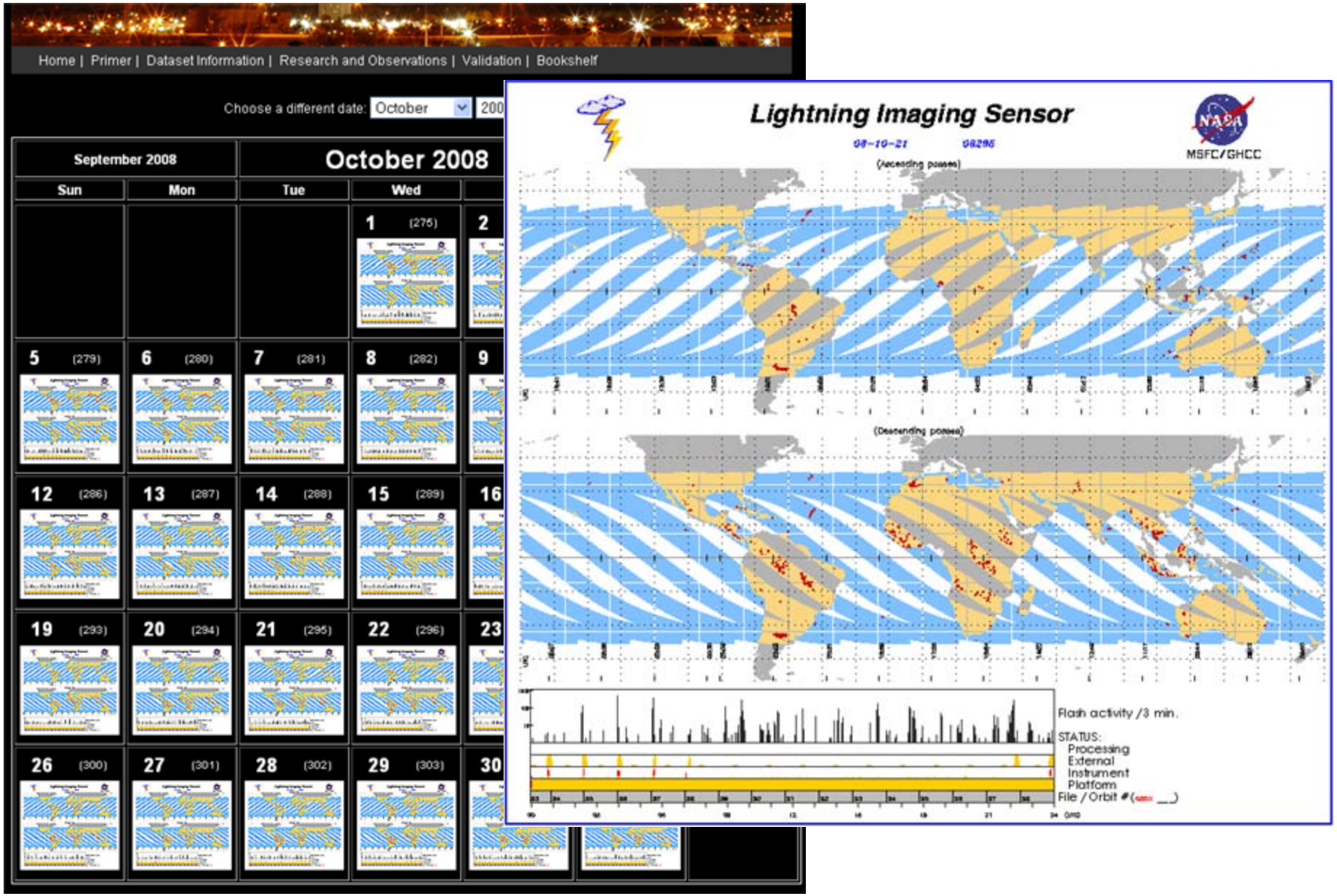


Advanced Microwave Precipitation Radiometer

Brightness Temperatures

Aircraft Based: NASA Field Campaigns
1993 - 2010

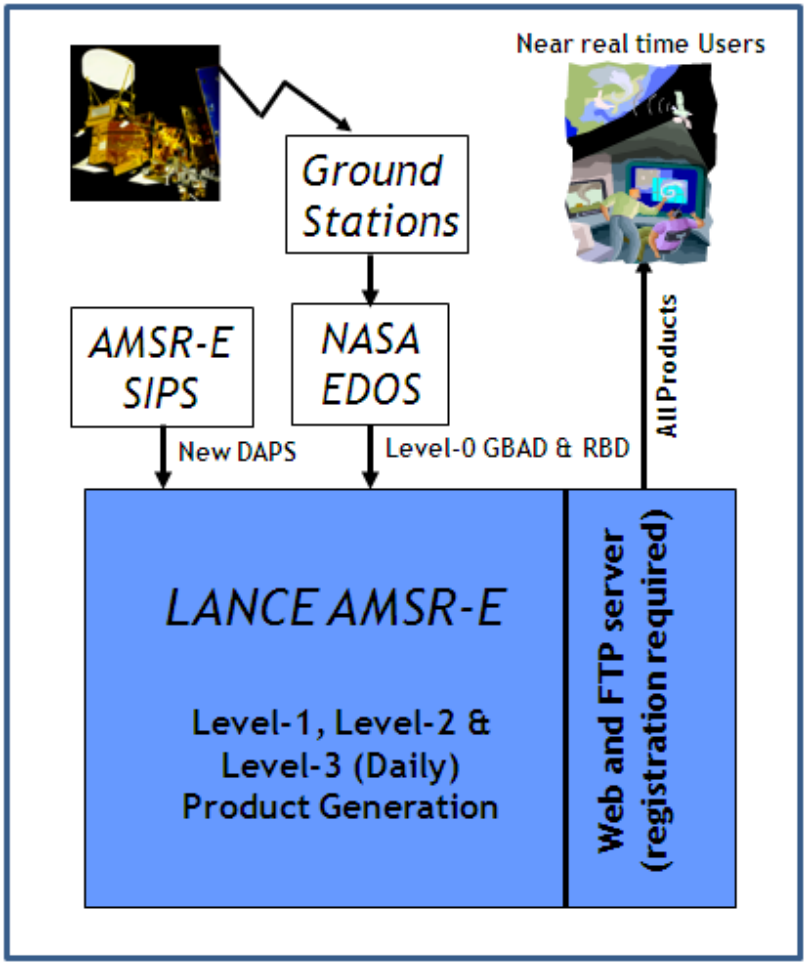
Twelve Year Global Lightning Data Record



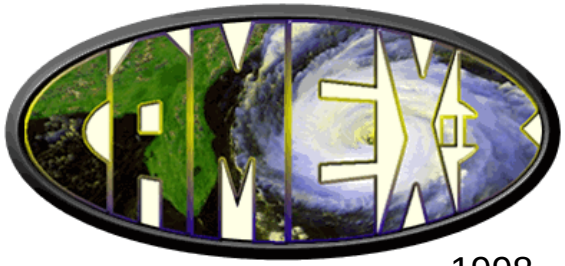
LANCE AMSR-E Near Real Time Data

Land Atmosphere Near-Real-Time Capability for EOS (LANCE) produces near-real-time data (<3 hours from observation) from Aqua AMSR-E. These products support a range of activities including weather forecasting and monitoring of natural hazards).

- Daily L3 12.5 km Polar Grids
  - Brightness Temperatures
  - Sea Ice Concentration
  - Snow Depth
- Daily L3 25 km Polar Grids
  - Brightness Temperatures
  - Sea Ice Concentration
- Daily L3 6.25 km Polar Grids
  - Brightness Temperature 89 GHz
- Daily L3
  - Global 0.25x0.25 deg Ocean Grids
  - Global Snow Water Equivalent EASE-Grids
  - Surface Soil Moisture EASE-Grids
- L2A Global Swath
  - Spatially-Resampled Brightness Temperatures
  - Ocean Products derived from Wentz Algorithm
- L2B
  - Global Swath Rain Rate/Type
  - Surface Soil Moisture



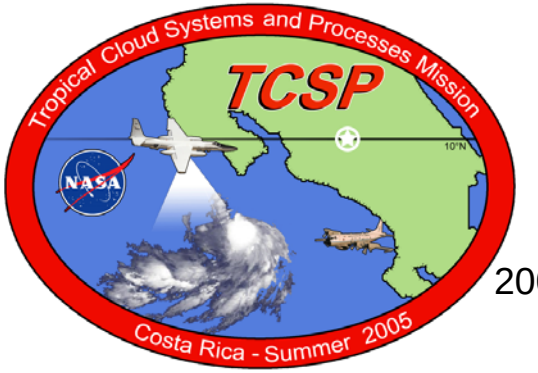
NASA Field Campaigns: 1998-2010



1998

2001

2002



2005

2006

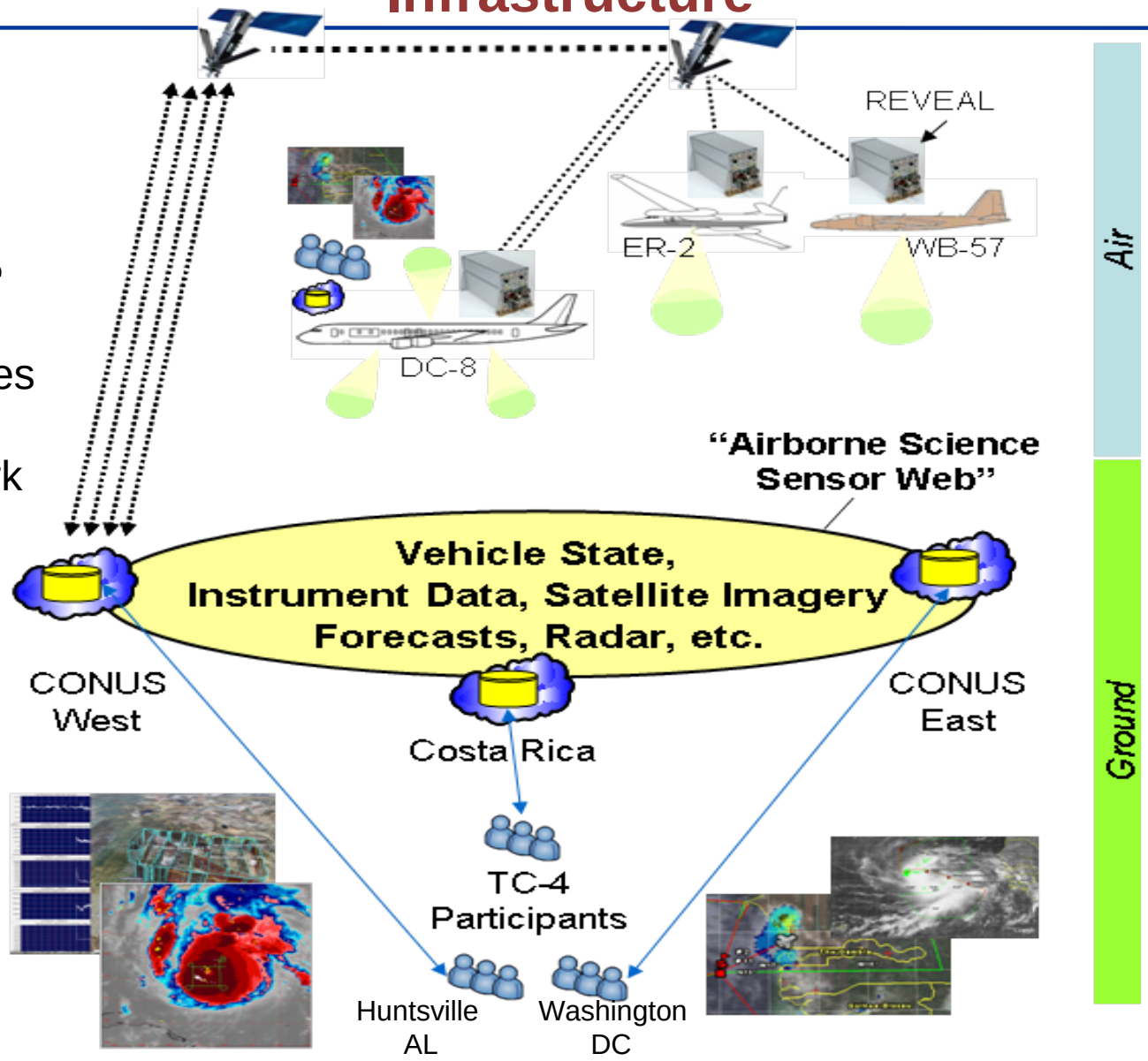
2010

2007

2008

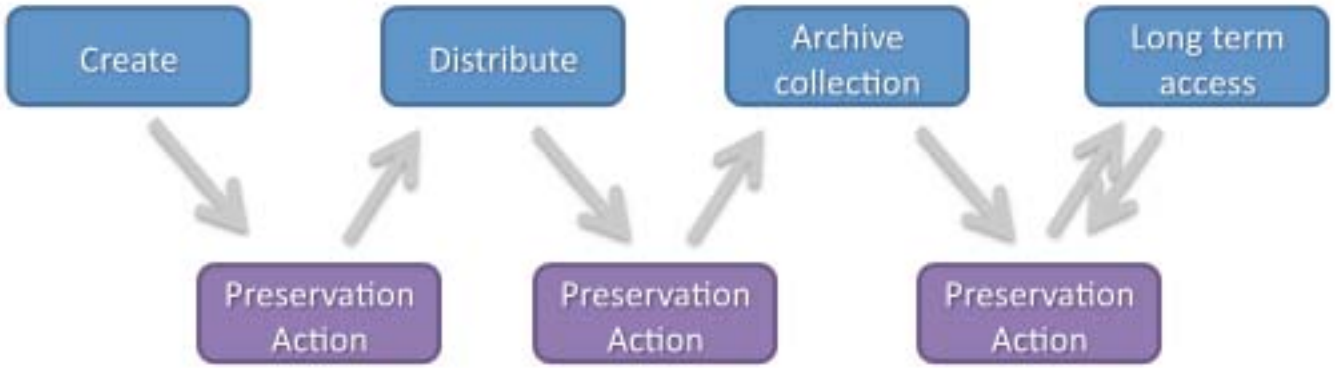
NASA Field Campaign Communications Infrastructure

REVEAL - *Research Environment for Vehicle Embedded Analysis on Linux* (REVEAL) provides satellite-based communication / network connectivity for aircraft



Key Responsibility at Data Creation

- Archivists are traditionally responsible for curation (i.e., adding metadata and provenance) but not present at creation of scientific data. Moment of creation is where most knowledge about product is present.



- To be effective, requires tools in earliest stages of data's life that help with preservation

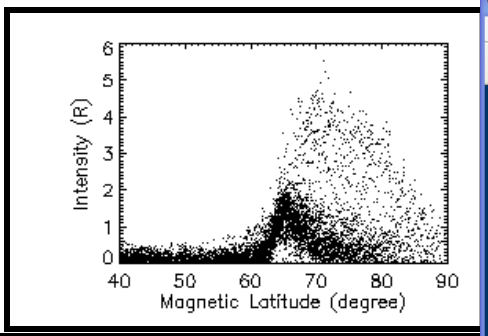
Diagram from Berman et al. "Sustainable Economics for a Digital Planet"

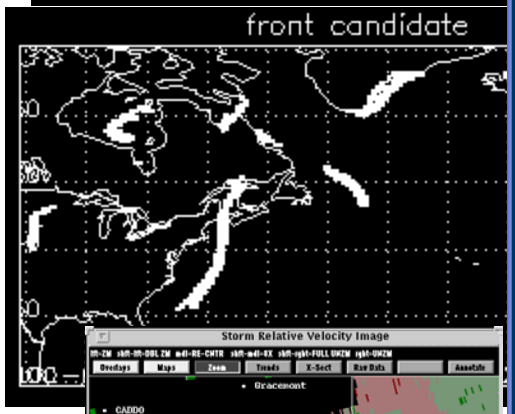
Types of Provenance Information

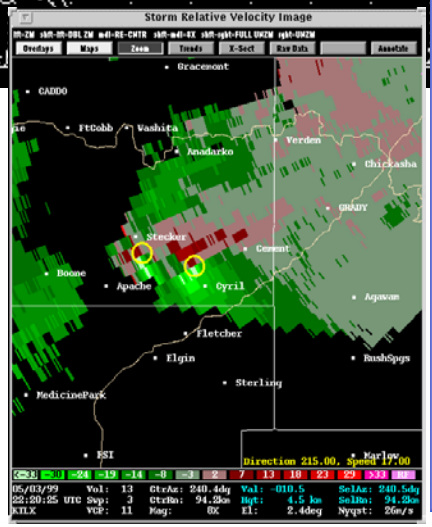
- ***Data lineage***: product generation "recipe" (data inputs, software and hardware)
- Additional ***knowledge*** about science algorithms, instrument variations, etc.
- ***Lots of information already available, but scattered across multiple locations***
  - Processing system configuration
  - Dataset and file level metadata
  - Processing history information
  - Quality assurance information
  - Software documentation (e.g., algorithm theoretical basis documents, release notes)
  - Data documentation (e.g., guide documents, README files)
- ***Instant Karma project aims to collate and organize information from multiple sources***

Data Mining: *Algorithm Development and Mining (ADaM) Toolkit*

- **UAHuntsville has been at the forefront of mining sensor data for over 15 years**
- **ADaM – UAHuntsville developed toolkit with 100+ algorithms**
- **Automated discovery of patterns, signatures, anomalies from vast observational data sets**
- **Derived knowledge for decision making and response**
- **Allows learning and training for adaptation**
- **Many users worldwide**





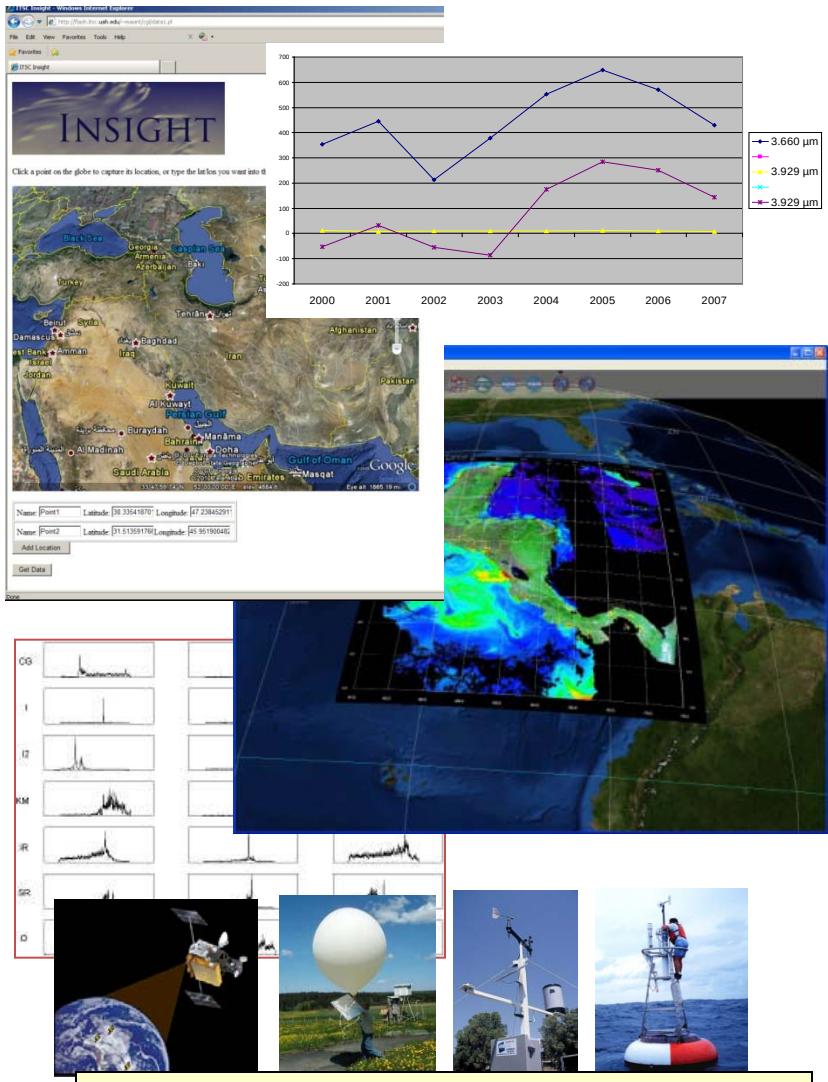


Data Mining: Environmental Data Analysis

How do you get the right information to the right people at the right time?

•Sensor Data Integration/Fusion
•Signature Analysis
•Pattern Recognition
•Real-Time Data Analysis

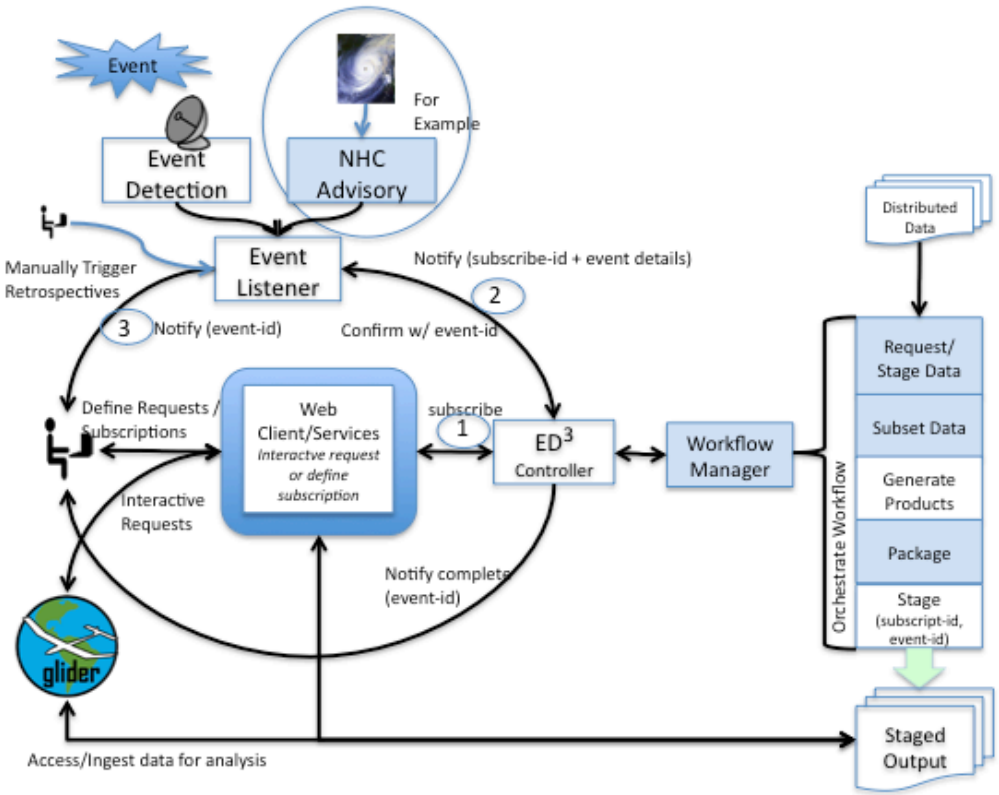
- **ADaM Algorithm Development and Mining toolkit**
- **MASINT Signature Analysis**
  - **Evaluating environmental impacts on national security**
- **NASA/USAID sponsored SERVIR Environmental Data Products for Central America and Africa**
  - **Decision Support System for environmental analysis**
- **JCTD EUCOM Efforts**
  - **Providing data products for the Arctic region**
- **NSF Linked Environments for Atmospheric Discovery**
  - **Real-time mining and analysis**
  - **Adaptive processing**



Sensor Data Integration is Critical for Adaptive Processing

Data Management and Dissemination for Analysis and Visualization

- ***Global Hydrology Resource Center:*** One of twelve NASA Earth science data centers; ingest, archive, processing and dissemination
- ***AMSR-E Science Investigator-led Processing System:*** Near-real-time and standard science data product processing and dissemination

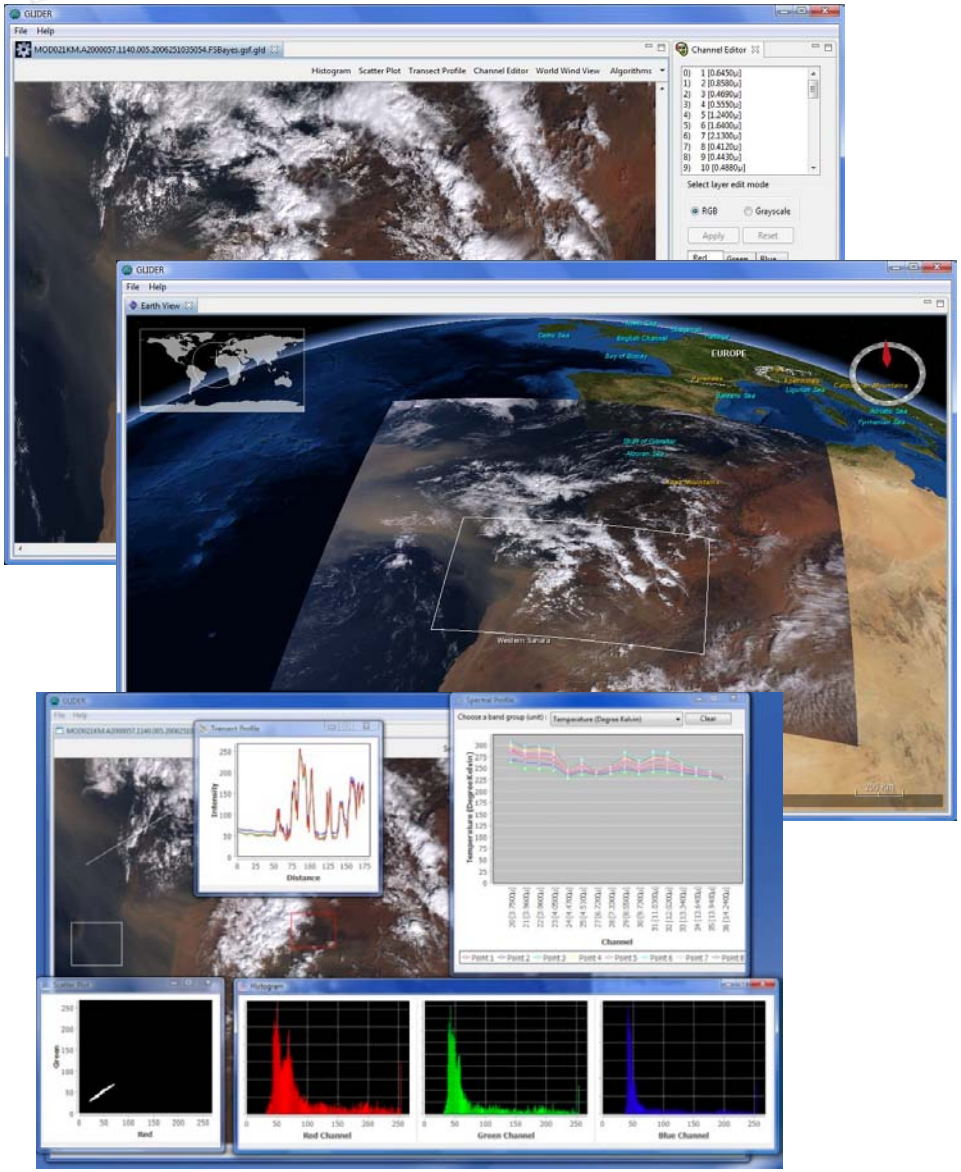


Event-Driven Data Delivery (ED3):

- ***Automated and discrete access*** to remote sensing data (NASA, NOAA, DOD, etc.)
- ***Event-Driven*** Data Delivery based on user inputs or subscriptions
- Enables ***adaptive processing***
- Can be integrated with GLIDER and other tools for ***mining, analysis, and visualization***
- Can be integrated with ***analysis workflow management*** tools

GLIDER

Globally Leveraged Integrated Data Explorer for Research



Capabilities:

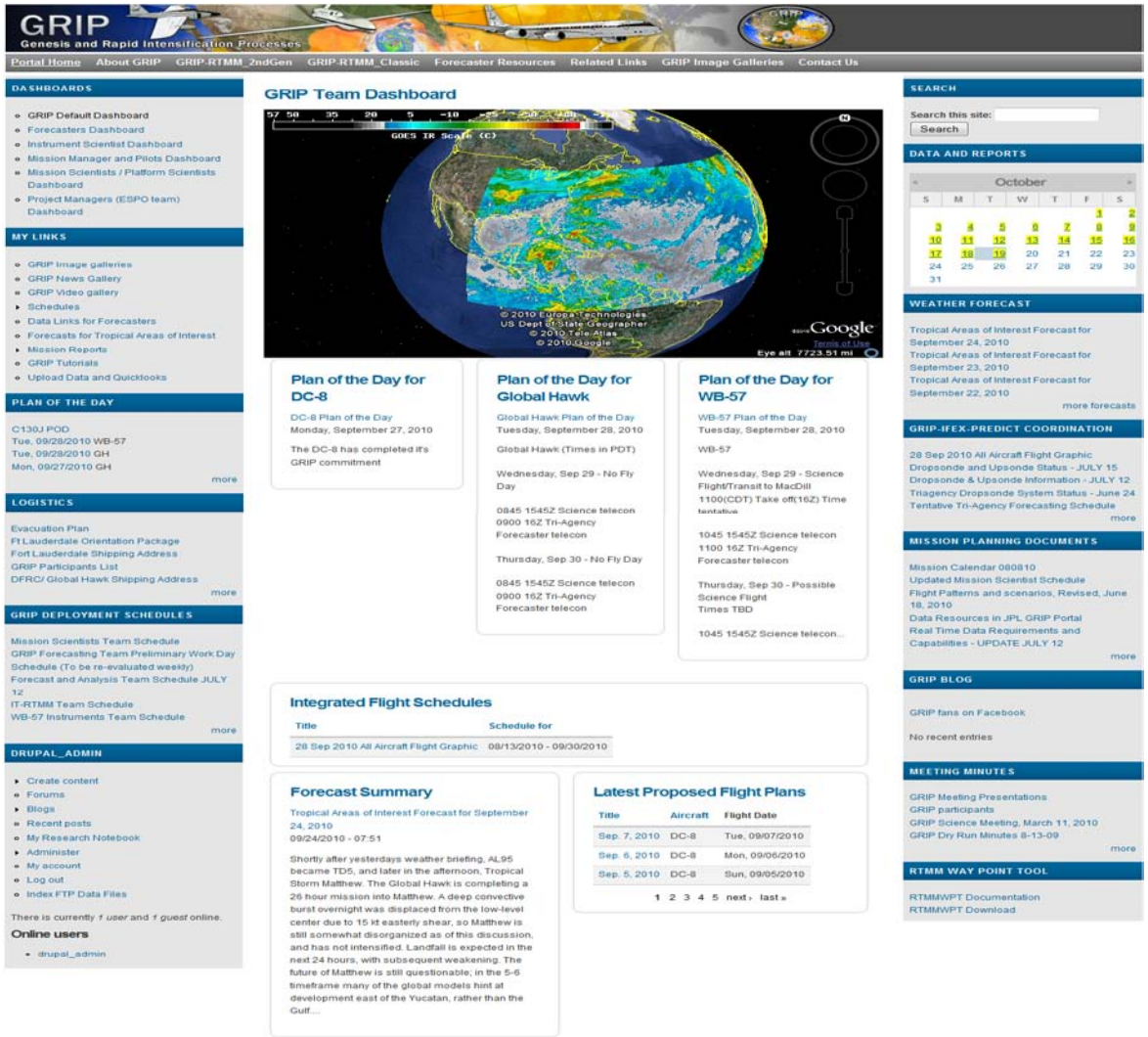
- ***Visualize and analyze*** satellite data in a native sensor view
- Apply ***image processing algorithms*** on the data
- Apply ***pattern recognition/data mining algorithms*** on the data
- ***3D Globe Visualization*** *of* satellite data, analysis/mining results, and additional layers
- Provides ***multiple views*** to manage, visualize, and analyze data

Integrates existing tools:

- ADaM: UAHuntsville's ***Algorithm Development and Mining*** Toolkit
- IVICS: UAHuntsville's ***Interactive Visualizer and Image Classifier for Satellites***
- WorldWind: NASA's ***3-D globe visualization*** system
- Other visualization tools

NASA Software Reuse Award 2010

Collaborative Environments:
Portals to Discovery

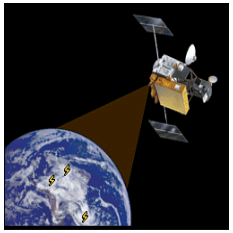


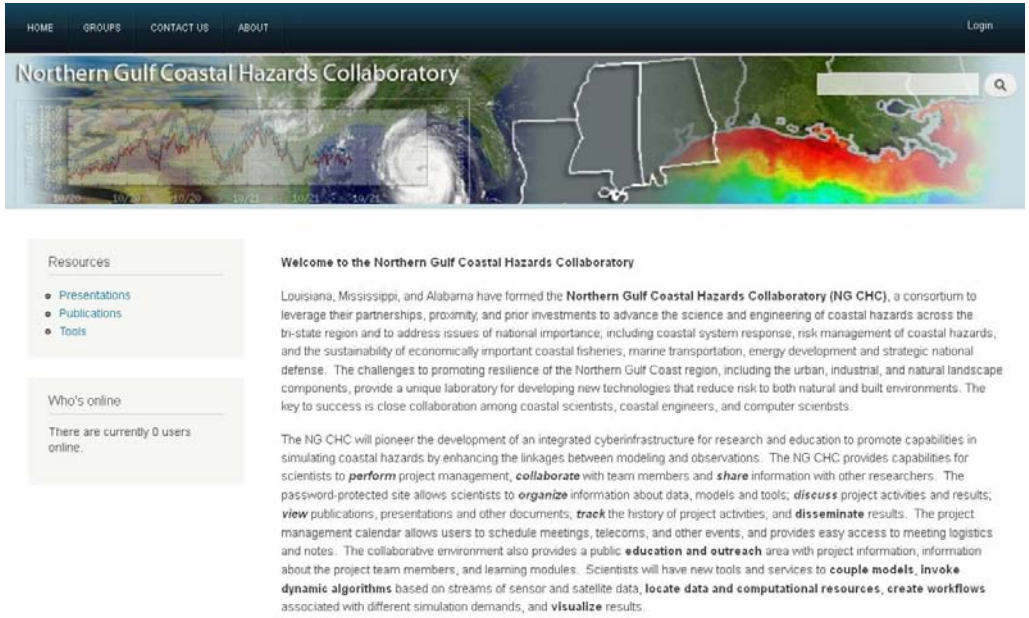
- Involve groups or teams to achieve common goals
- share ideas, documents, media, tools
- Users can submit or update contents
- Improve communication
- Control access
- Manage and index contents
- Categorize contents
- Integrate semantics

Northern Gulf Coastal Hazards Collaboratory

Research and Education Cyberinfrastructure Investments to Develop the Coastal Hazards Collaboratory in the Northern Gulf Coast

a National Science Foundation (NSF) Experimental Program to Stimulate Competitive Research (EPSCoR) Research Infrastructure Improvement (RII) Track-2 program





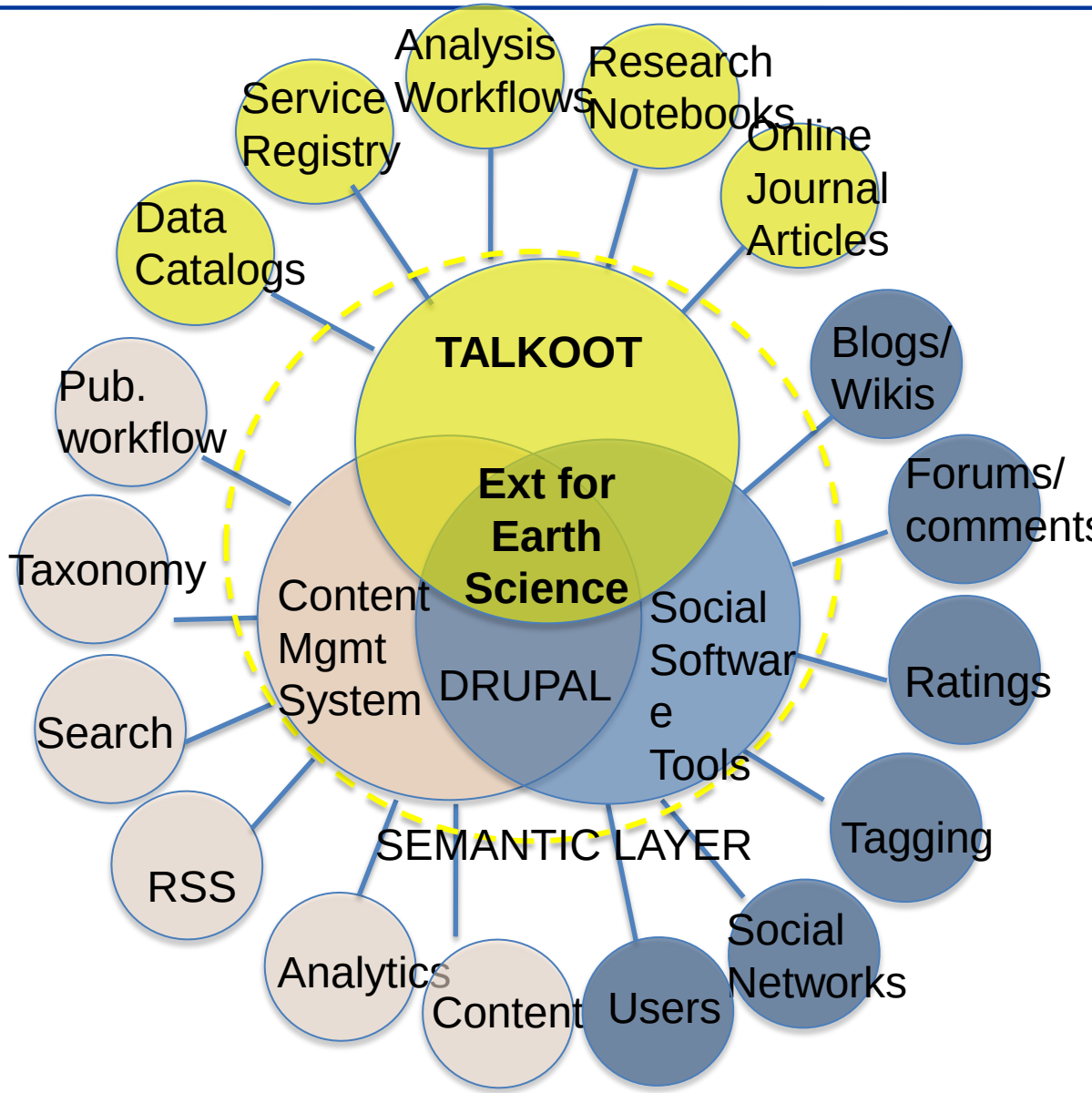
Promoting coastal resiliency with new technologies that reduce risks to both natural and built environments

Addressing engineering design, coastal system response, and risk management to coastal hazards

Catalyzing collaborative research via advanced cyberinfrastructure

Portal Technology: Talkoot

- Customizable "software appliance" to build collaborative portals
- Leverages Drupal, an open architecture platform, to provide the core Content Management System capabilities
  - Drupal also has a vast array of specialized modules that have been developed by its user community to provide additional features
- Adds Earth science specific modules to provide data searching, processing , analysis, visualization plus many other capabilities



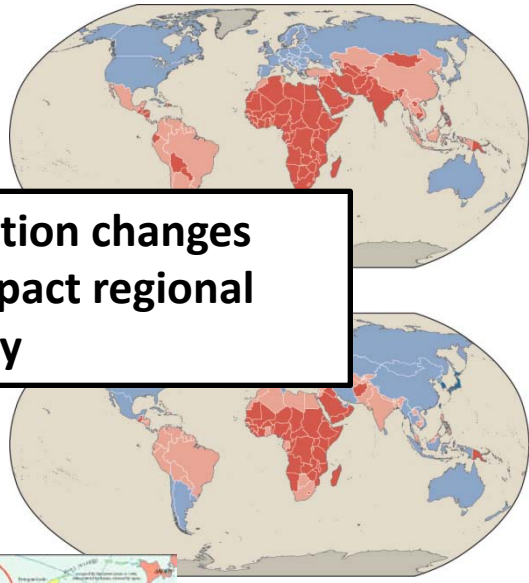
Environmental Impacts on National Security

- Provide information for use in planning and understanding the impact of environmental issues on national security
- Analysis of Environmental and Climate Impacts
  - Identify possible security risks due to the interaction of environmental conditions, climate impacts and human activity
  - Use historical datasets for trends and patterns
  - integrate these results with:
    - current environmental measurements
    - signatures of environmental change (e.g. sea ice and land use)

- Initially focus on the relationships among land use change and moisture conditions in Afghanistan and sea ice in the Arctic
  - Identify potential future problem areas
  - Extend and adapt to address other environmental security issues
  - Focus on the needs of both military and intelligence communities
  - Analyze contributing factors toward instability in regions of interest

National Intelligence Council "Global Trends 2025" and "Global Governance 2025"

The multiple links among climate change and resources issues; the economic crisis; and state fragility—"hubs" of risks for the future—illustrate the ***interconnected nature*** *of the challenges on the international agenda today. –Global Governance 2025*



Population changes will impact regional stability

- NIC: Environmental conditions and man's activities will impact the stability of nations around the world
  - Food: Demand will rise by as much as 50 percent, putting pressure on many nations
  - Water: Access to fresh water will become increasingly difficult for some
  - Population Changes: Some nations will see major growth, others major declines
  - Climate Change: Impact will vary by region – we need to model the possibilities
  - New Technologies: May impact land use and energy production
  - Energy Transformation: New power sources may lead to changes in the use of natural resources

The opening of the Arctic will impact trade and shift power

http://www.dni.gov/nic/NIC_2025_project.html

National Intelligence Council

We continue to assess that global climate change will have wide-ranging implications for US national security interests over the next 20 years because it will aggravate existing world problems—such as poverty, social tensions, environmental degradation, ineffectual leadership, and weak political institutions—that threaten state stability.

As climate changes spur more humanitarian emergencies, the demand may significantly tax US military transportation and support force structures, resulting in a strained readiness posture and decreased strategic depth for combat operations.

The effects of climate change in North Africa are likely to exacerbate existing threats to the region's water and food resources, economies, urban infrastructure, and sociopolitical systems

Climatic stress coupled with socioeconomic crises and ineffective state responses could generate localized social or governmental collapses and humanitarian crises.

Arctic states have such common goals as environmental protection, shipping safety, effective search and rescue (SAR), and commerce development, they do not fully agree on how to achieve them. Our research indicates the widespread use of the Arctic for commercial purpose is hindered by the absence of reasonable insurance, which will come only with Arctic coastal infrastructure development, agreed safety standards for commercial vessels, and adequate SAR capability.

- Dennis C. Blair's Congressional Testimony, 2010

NIC's Rich Engel's Challenges from the Scripp's Climate Change and National Security Symposium – June 2010

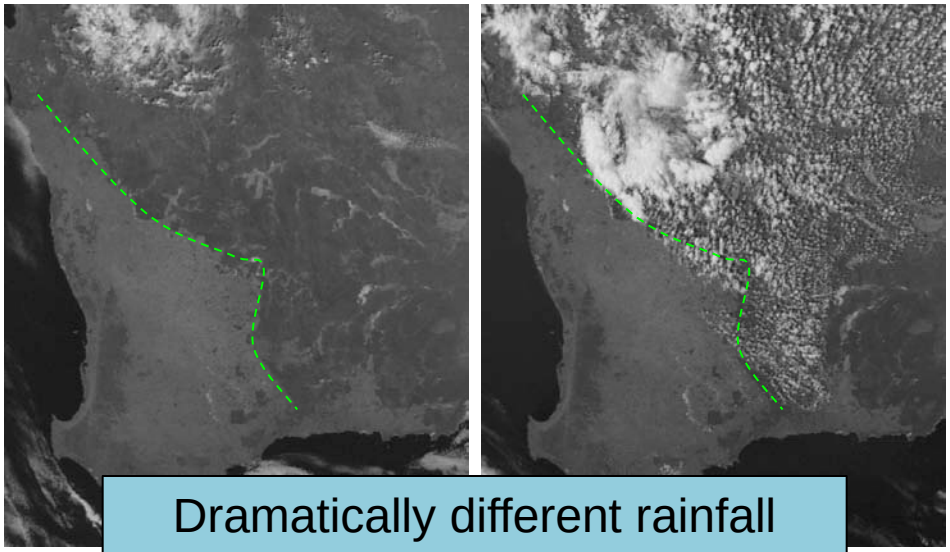
- **Global climate change will have wide-ranging implications for US national security interests over the next 20 years**

- **Changes in water availability - people move**
- **Changes in agriculture productivity – people move**
- **Damages to economically significant infrastructure from extreme weather events.**
- **Change in disease patterns (human, plant, animal)**

- Need better Arctic Situational Awareness
- Ice sheets and movements
- Weather
- Shipping Traffic

http://www.dni.gov/testimonies/20090212_testimony.pdf

Impact of human activity on rainfall



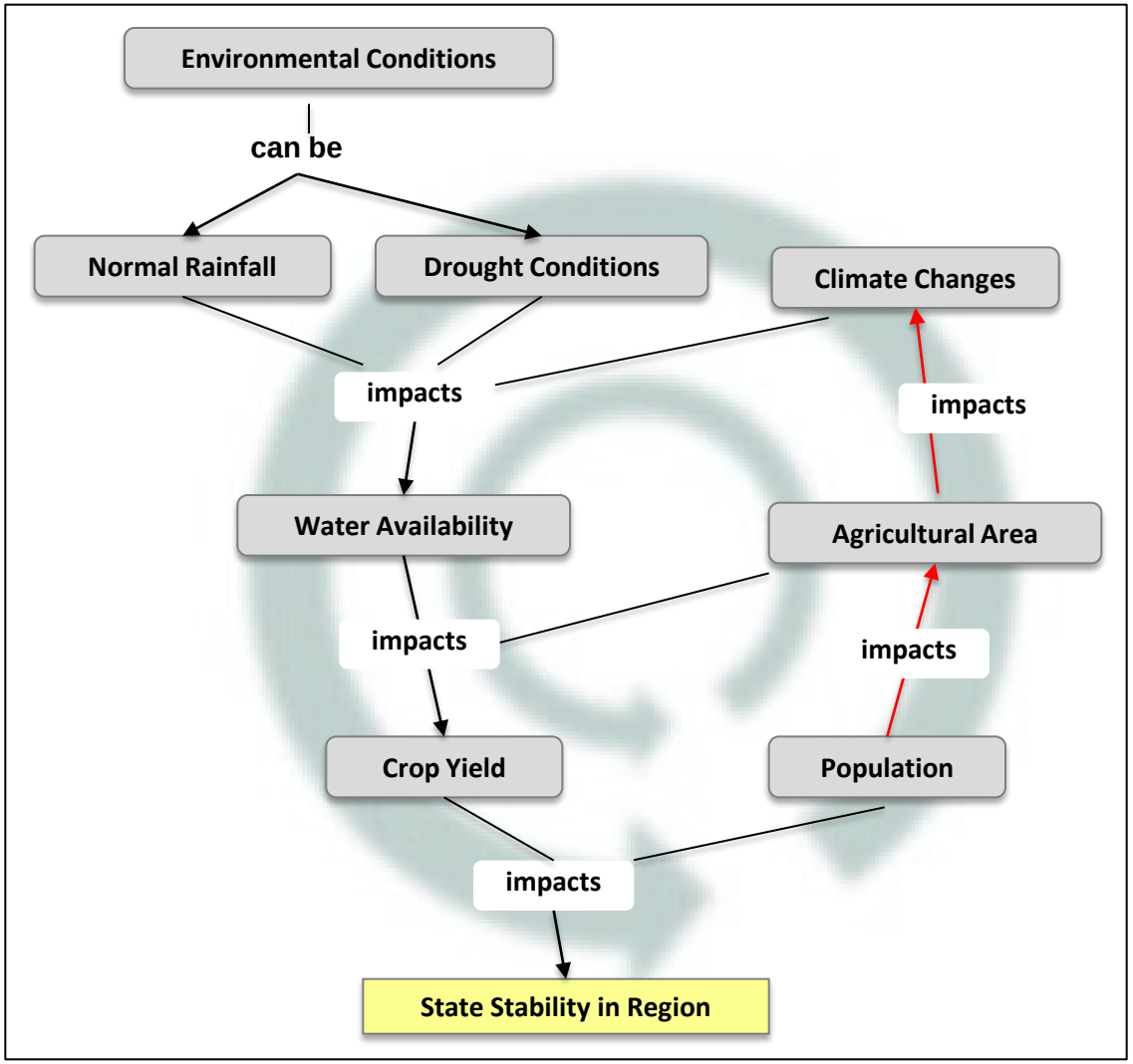
Dramatically different rainfall amounts due to human activity

- Land clearing in Australia has led to changes in rainfall and weather patterns
- The impact on agriculture can be substantial

These changes have led to salinization and irreversible cropland destruction

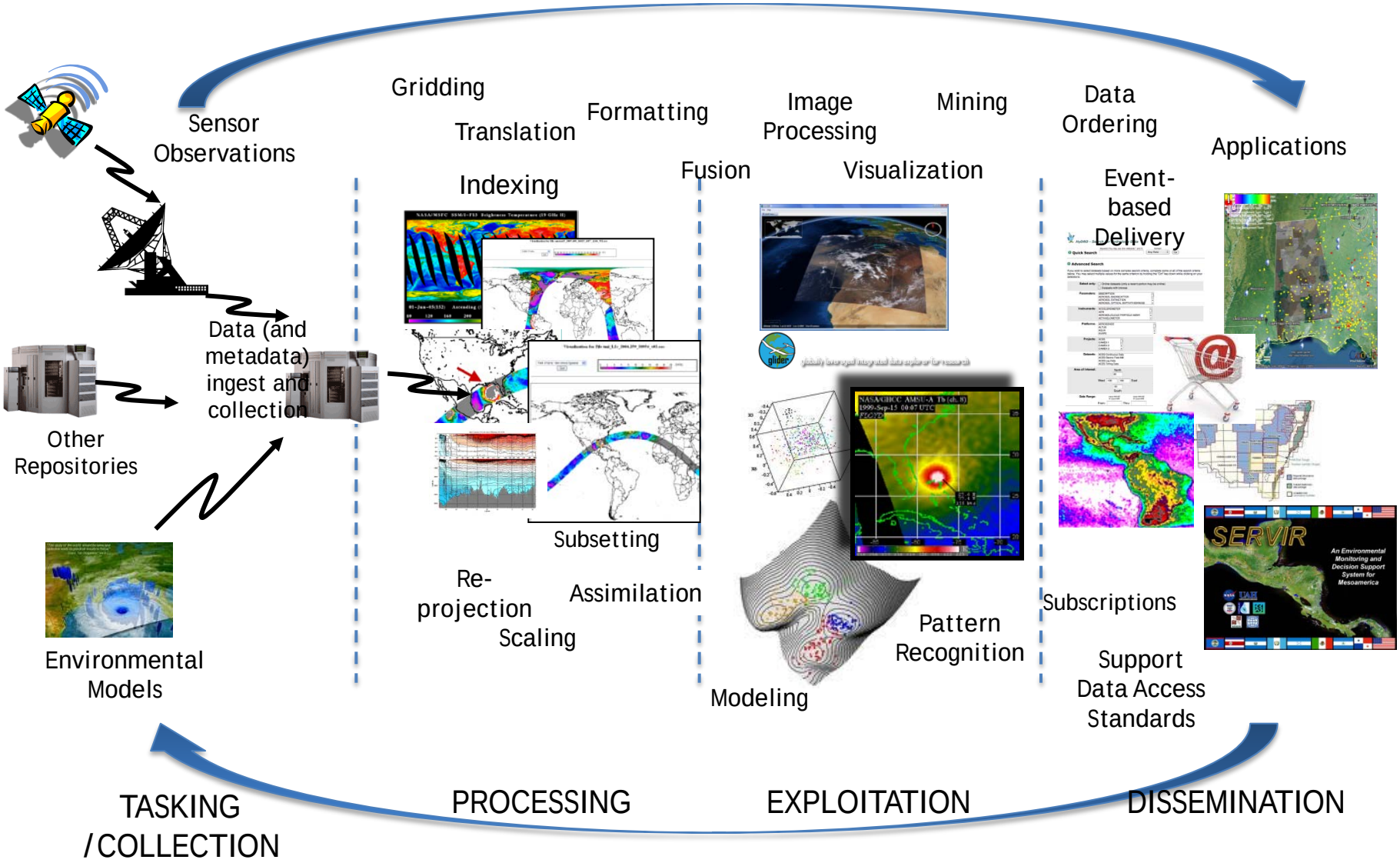
Coupled Human-Environment System

This effort will develop a methodology to analyze the coupled human-environment system in order to understand how human utilization of land influences regional climate, impacting in turn agricultural productivity and human sustenance.



Unlike global climate change impacts, mitigation of regional impacts is more tractable and could be effected through external aid.

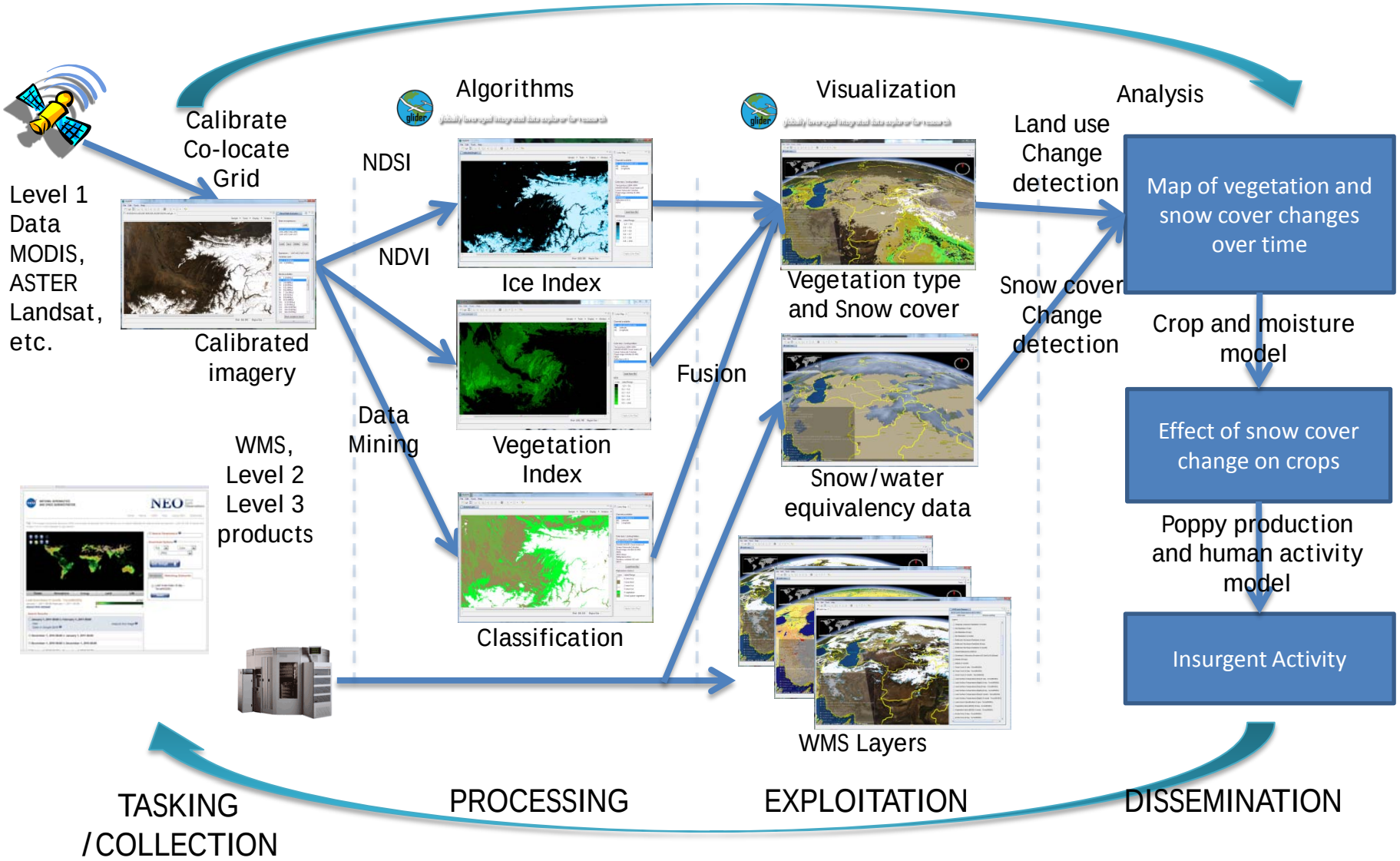
Data Integration and Interoperability Framework



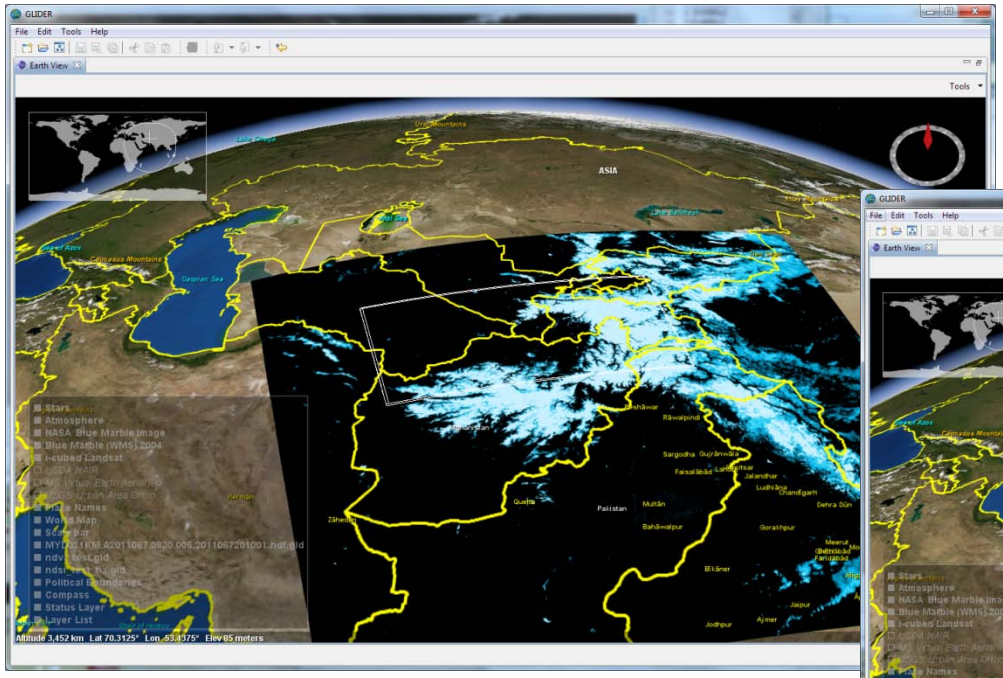
Example: Mountain Snow Cover In Afghanistan

- Coupled System Hypothesis: Winter mountain snow cover and poppy crop production
- Assumed human impact: Increase in snow cover leads to increased poppy and food crop yields
  - Poppy crop production funds insurgent activity
  - Increase in poppy crop leads to increased insurgent activity
  - Decrease in food crops causes instability, aiding insurgents
- Possible mitigation: Analysis of possible crop yields may allow planning for counter insurgent activity

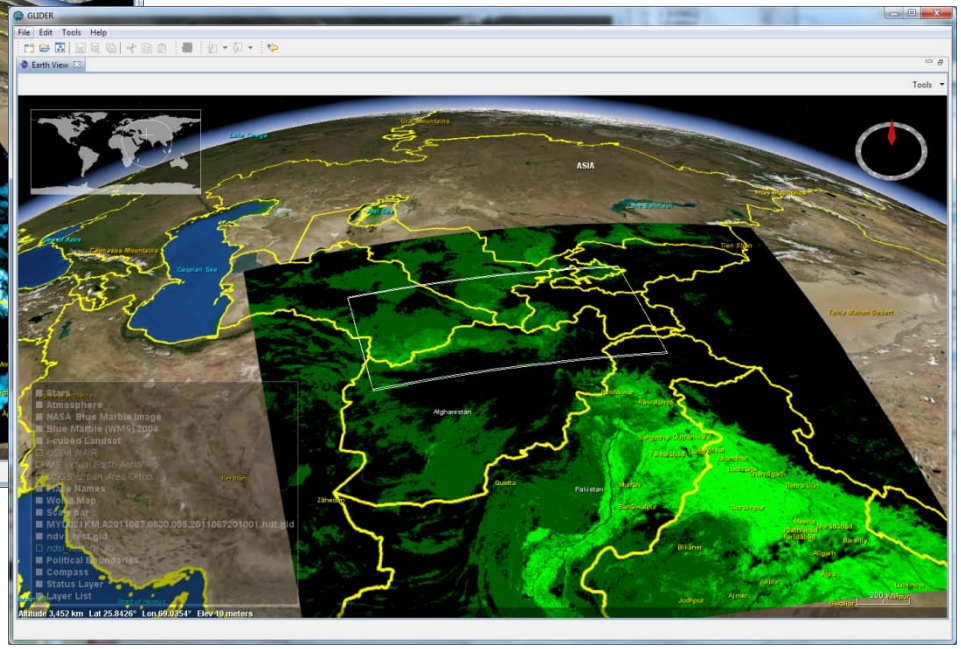
Mountain Snow Cover Process



Utilizing Glider to Analyze Relationship Between Snowfall and Agricultural Activity



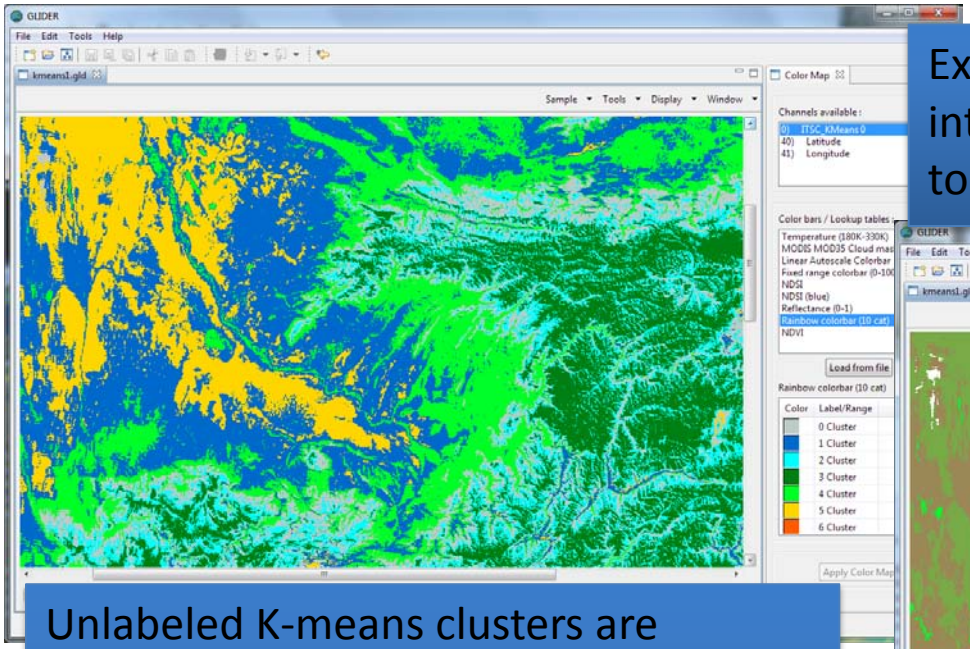
Snowcover and vegetation in Afghanistan derived using MODIS data.



These two products can be used to study how water is linked to poppy and wheat production.

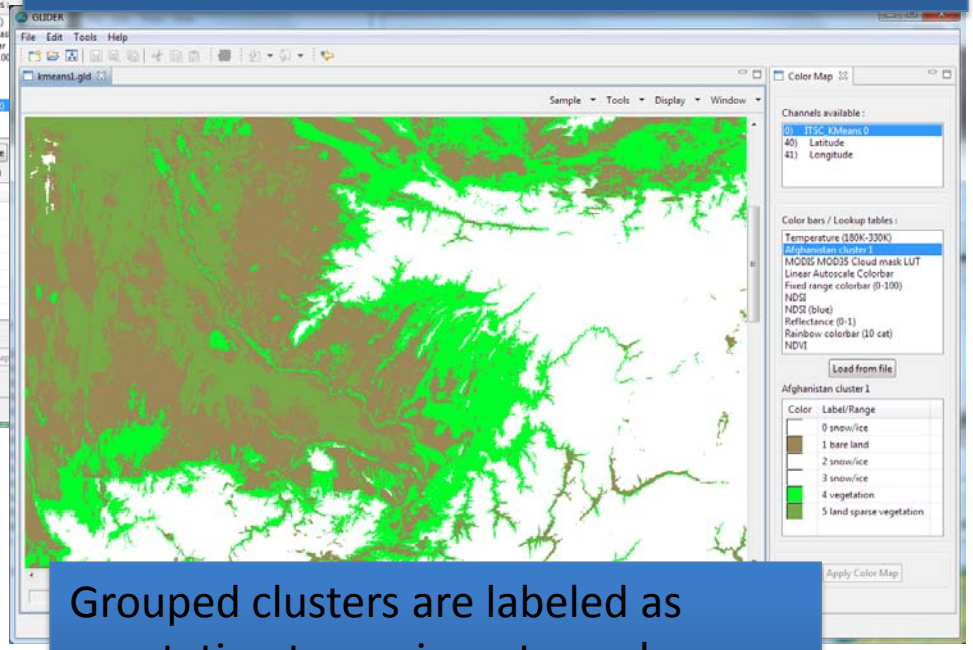
Thematic Map of Surface Types Using Glider

- Clustering algorithms may be used to provide the classification capabilities required to determine crop type in Afghanistan



Example using K-means to cluster regions of interest together. Glider is providing the basic tools for this analysis.

Unlabeled K-means clusters are displayed as different colors. These clusters can be grouped and labeled as different land/surface types.



Grouped clusters are labeled as vegetation types, ice, etc. and displayed as specified colors.

Collaborations, Connections, and Publications

- NGA, NASIC
- EUCOM, NORTHCOM
- DHS, DOE
- NASA, NOAA
- SMDC, MSIC, AMCOM, MDA
- International Partners

Publications:

Cloud Cover Comparisons of the MODIS Daytime Cloud Mask With Surface Instruments at the North Slope of Alaska ARM Site, T.A. Berendes, D.A.; Berendes; R.M. Welch, E. G. Dutton; T. Uttal; E. E. Clothiaux, IEEE Trans Geosci Remote Sens, Vol. 42, No. 11, pp. 2584- 2593, Nov. 2004.

Validation of a neural network based MODIS global cloud mask using ground-based instruments, Todd Berendes, D. A. Berendes, R. M. Welch, E. E. Clothiaux, E. G. Dutton, P. Minnett, T. Uttal, AMS 18th International Conference on Interactive Information and Processing Systems for Meteorlogy, Oceanography, and Hydrology, January 14, 2002, Orlando, FL.

The Aster Polar Cloud Mask, A. M. Logar, D. E. Lloyd, M. L. Penaloza, R. E. Feind, T. A. Berendes, K. S. Kuo, R. M. Welch, IEEE Trans Geosci Remote Sens, 36, 1302-1312, 1998.

Clouds and the Earth's Radiant Energy System (CERES) Algorithm Theoretical Basis Document Subsystem 4.1 – Cloud Mask, Bryan A. Baum, Ronald M. Welch, Pat Minnis, Larry L. Stowe, James A. Coakley, Jr., Qing Trepte, Pat Heck, Xiquan Dong, Dave Doelling, Sunny Sun-Mack, Tim Murray, Todd Berendes, Kwo-Sen Kuo, Paul Davis, Official document of the EOS CERES Science Team. 1997.

Impact of land use on tropical montane cloud forests: 1. Sensitivity of cumulus cloud field characteristics to lowland deforestation, Nair U. S., R. O. Lawton, R. M. Welch, R. A. Pielke.Sr., 2003:, J. Geophys. Res., 108, NO. D7, 4206, doi:10.1029/2001JD001135.

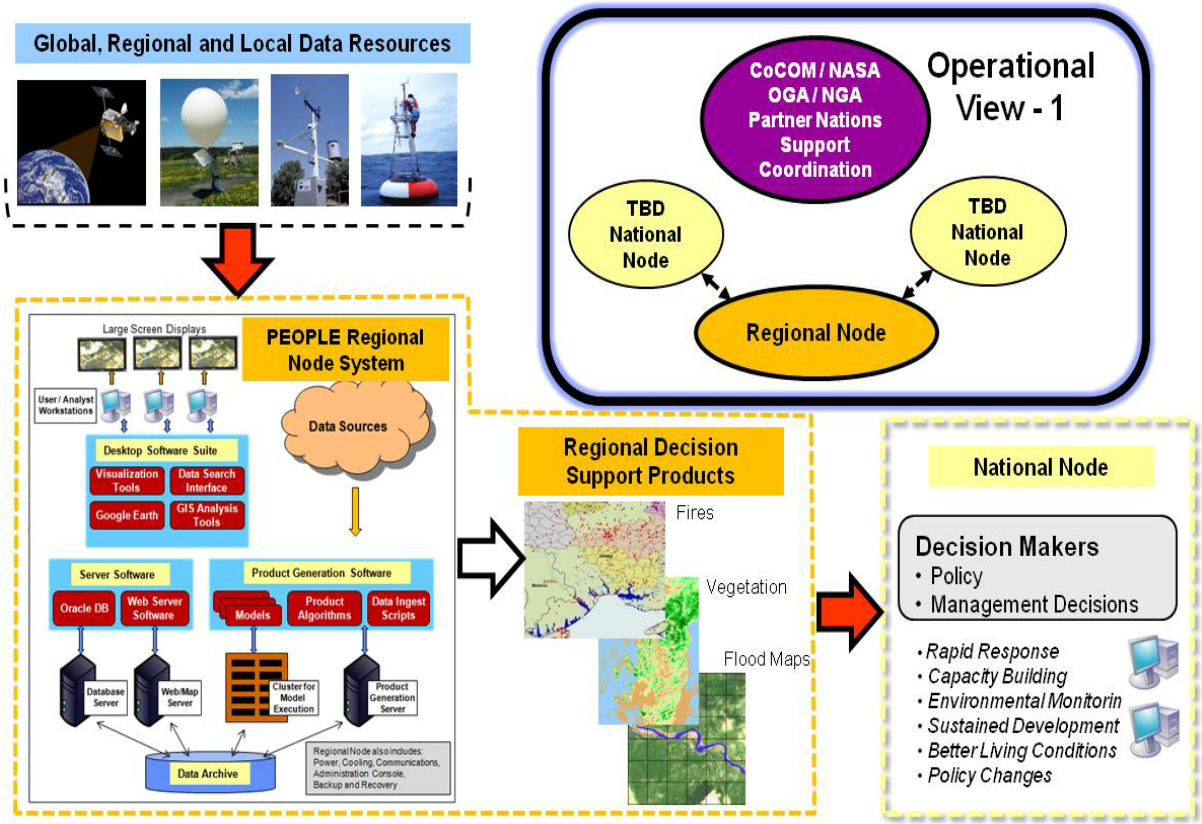
PEOPLE JCTD: Partnering Earth Observations for People Living Environmentally Joint Capability Technology Demonstration

- Deliver a low cost operational system with processing power, analysis software and extensive data resources
- Capitalize on proven government, university and commercially-owned technologies and readily available datasets, tools, and products

One-Year Deliverable

- Operational integrated architecture for Regional Node and two or more Partner Nation nodes
- Product generation software, operational models, visualization and analysis software, and integrated data sets ready for transition

Provide access to both sensor data and derived data products for countries and regions of military interest – focus is on nation building and environmental impacts



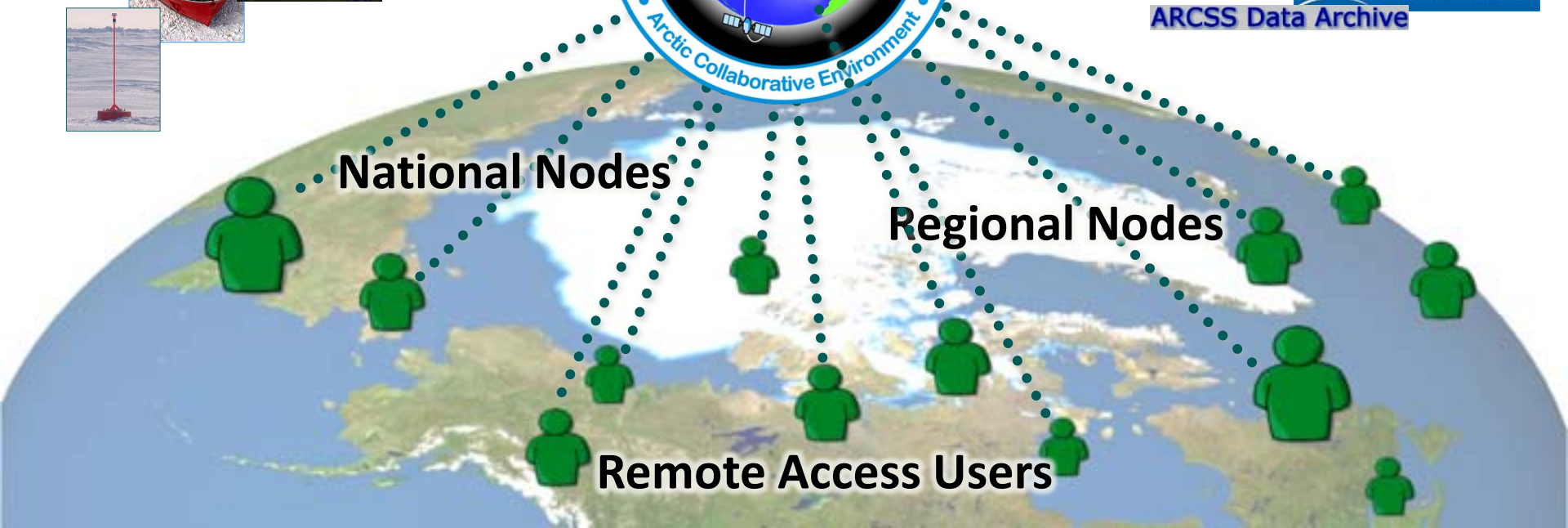
Links with other efforts: PEOPLE-ACE JCTD

An open source, web-based, multi-national environmental monitoring, research, and decisions support system to enable development of advanced value-added products

International observing systems & product developers

etc.

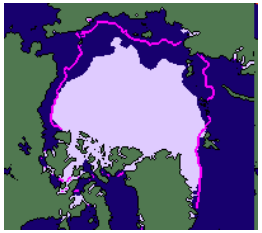
Data portals and science programs



Example: Improved Sea Ice Analysis

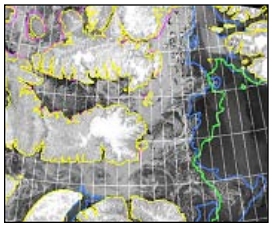
Observation and Analysis Inputs

Passive microwave observations



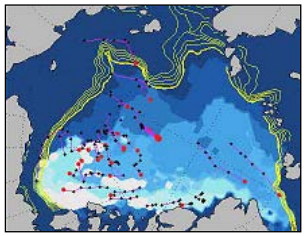
Visible observations including declassified imagery

Radar & LiDAR sea ice mapping

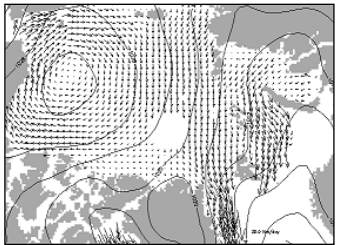


Ice thickness charting

Buoy drift and ice concentration



Ice motion by remote sensing & modeling



Data Integration, Assimilation & Analysis

Products for Decision Support

- Verification of sea ice extent, area
- Improved freeboard and thickness estimates
- Improved ice age estimates
- Improved ice drift modeling
- Improved spring melt and fall accretion modeling
- Reduced errors for observed physical states
- Improved prediction of ice edge location
- Improved prediction of above states with reduced uncertainty
- Improved understanding of sea ice effects on polar energy balance

PEOPLE-ACE Priorities

Missions

- Maritime SAR
- On-Ice SAR
- Maritime Safety
- Humanitarian Response
- Environmental Response
- In-Situ Mobility
- Scientific Research
- Infrastructure Design & Development
- Inter-Agency/International Cooperation
- Inter-Agency/International Training

Tasks

- **Plan and Execute Maritime Navigation**
- **Data & Voice Communications**
- **Avoid Superstructure Icing**
- **Share Arctic Data**
- **Locate and Assess Surface Ships**
- **Assess Snow & Ice Coverage**
- **Assess Tides & Currents**
- **Mitigate Infrastructure Icing**
- **Develop & Share Constructed Data**
- **Assess Permafrost Conditions**
- **Avoid Airfoil and Structural Icing**

Products

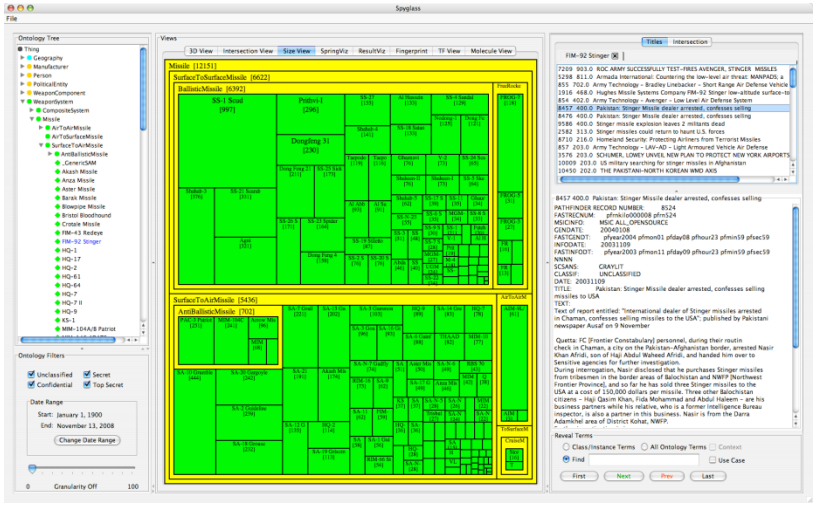
- **Sea-Ice Onset, Growth, Drift, and Decay**
- **Sea Surface Temperature**
- **Air Temperature**
- **Air Pressure & Gradients**
- **Wind Speed, Direction, Gusts, & Shear**
- **Tides, Currents, and Swells**
- **Relative Humidity & Dew Point**
- **Weather Forecasting**
- **Indigenous Weather Knowledge**
- **Propensity for Superstructure Icing**
- **Changes in Shoreline**
- **Soil Moisture**
- **Vegetation**
- **Surface Roughness**
- **Vessel Location and Characterization**

SpyGlass
Ontological Text Mining

Multiple Visualization Approaches

- Show distribution of documents across categories
- Show relationships between documents / categories

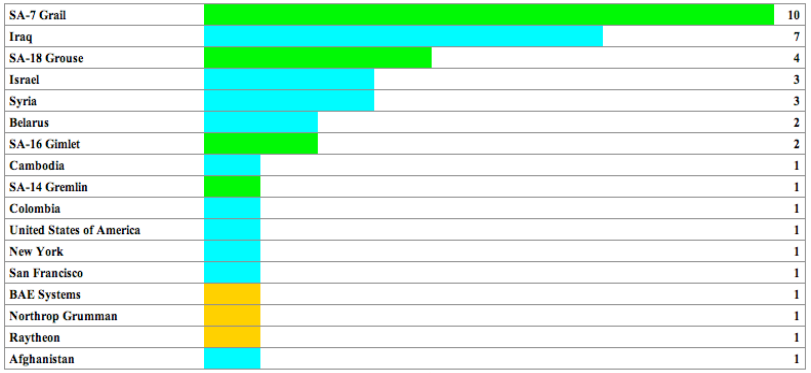
Ontology



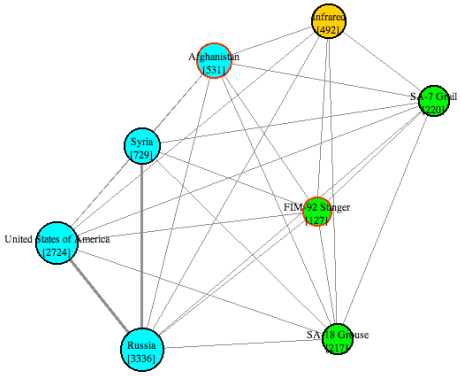
Document Browse View

Size View

- Captures ***semantic information*** and ***contextual knowledge*** of analysts
- Ontology describes ***entities, concepts and relationships*** in a domain
- ***Constructs document index*** for each term, listing all documents where term occurs
- ***Fast indexing and retrieval***, with high precision and recall
- ***Scores documents*** by number of relevant terms
- ***More powerful*** than simple keyword queries
- ***Possible to reason*** over ontological structures



Document Fingerprint View



Molecule View of Concept Associations

http://www.itsc.uah.edu