



Workshop on Migration of Agriculture as One Path to Sustainability

Sara J. Graves

Director, Information Technology and Systems Center

UA System Board of Trustees University Professor

Professor, Computer Science Department

256-824-6064

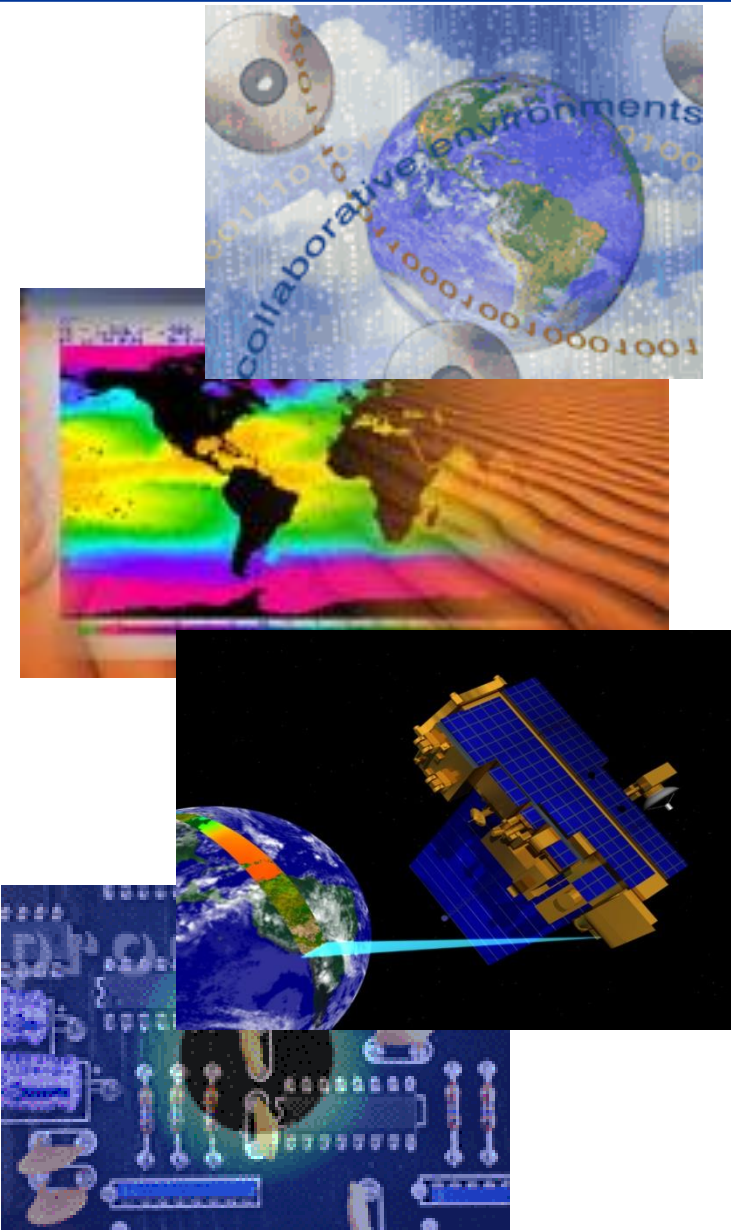
sgraves@itsc.uah.edu

Boulder, CO

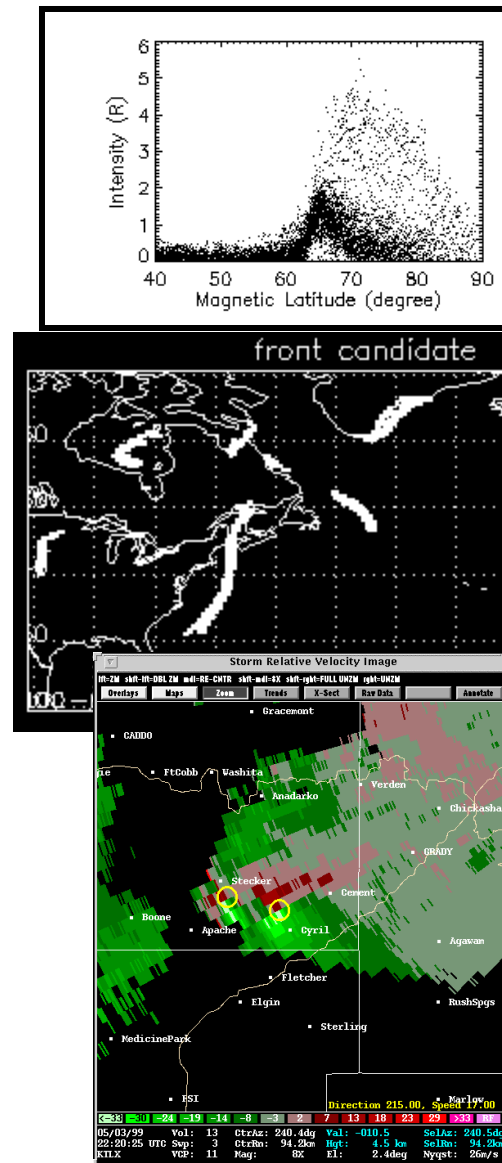
October 21-23, 2015

Primary Research Focus Areas

- Data Mining and Knowledge Discovery
- Visual Analytics
- Geoinformatics
- Modeling and Simulation
- Systems Interoperability
- Knowledge Engineering and Semantic Web
- Information Management and Technologies for Big Data
- Asynchronous Collaboration Technologies / Web 2.0
- Decision Support Frameworks
- Cybersecurity/Information Assurance
- Real-time Processing
- Event-driven / On-demand Processing
- Geospatial Technologies
- Sensor Networking Analysis



- UAHuntsville has been at the forefront of mining sensor data for over 15 years
- ADaM – UAHuntsville developed toolkit with 100+ algorithms, used worldwide
- Automated discovery of patterns, signatures, anomalies
- Derived knowledge for decision making and response
- Allows learning and training for adaptation
- Most cited article in *Elsevier Computers and Geosciences*, 2005-2010



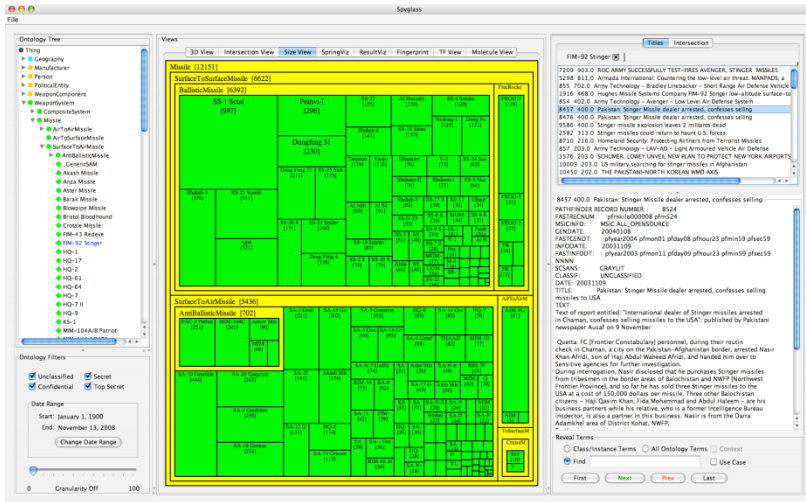

The screenshot shows a Mozilla Firefox browser window displaying the ADaM website. The browser title is 'ADaM - Mozilla Firefox' and the address bar shows 'http://datamining.itsc.uah.edu:3945/adam/doci'. The website header includes 'THE UNIVERSITY OF ALABAMA IN HUNTSVILLE INFORMATION TECHNOLOGY AND SYSTEMS CENTER' and 'ADaM Algorithm Development and Mining system'. The main content area is divided into several sections: 'Home News Publications To Do Bug Tracker', 'Get Involved Mailing Lists', 'Download License Binaries', 'ADaM Help Documentation FAQ Links', and 'Contact Us'. There are also links for 'ADaM Documentation', 'Data Mining Overview', and 'ADaM 4.0.2 Overview'. A section titled 'ADaM 4.0.2 Components' lists 'Pattern Recognition' and 'Image Processing' techniques. The 'Pattern Recognition' section includes 'Classification Techniques' (Bayes Classifier, Naive Bayes Classifier, Bayes Network Classifier, CBEA Classifier, Decision Tree Classifier, SEA classifier, Very Fast Decision Tree Classifier, Back Propagation Neural Network, k-Nearest Neighbor Classifier, Multiple Prototype Minimum Distance Classifier, Recursively Splitting Neural Network) and 'Clustering Techniques' (DBSCAN, Hierarchical Clustering, Isodata, k-Means, k-Medoids, Maximin). The 'Image Processing' section includes 'Basic Image Operations' (Arithmetic Operations(+/-), Collaging, Cropping, Image Difference, Image Normalization, Image Moments, Equalization, Inverse, Quantization, Relative Level Quantization, Resampling, Rotation, Scaling, Statistics, Thresholding, Vector Plot) and 'Segmentation/Edge and Shape Detection' (Boundary Detection, Polygon Circumscription, Making Region, Marking Region). There is also a 'Filtering' section with 'Dilation'.



analyze and search large text sources (databases, on-line resources, etc.)

Multiple Visualization Approaches

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



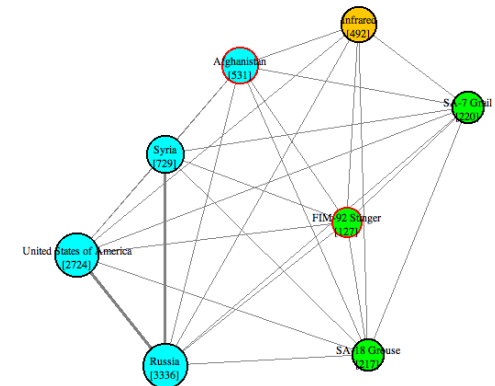
Size View

SA-7 Grail	10
Iraq	7
SA-18 Grouse	4
Israel	3
Syria	3
Belarus	2
SA-16 Gimlet	2
Cambodia	1
SA-14 Gremlin	1
Colombia	1
United States of America	1
New York	1
San Francisco	1
BAE Systems	1
Northrop Grumman	1
Raytheon	1
Afghanistan	1

Document Fingerprint View

Document Browse 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

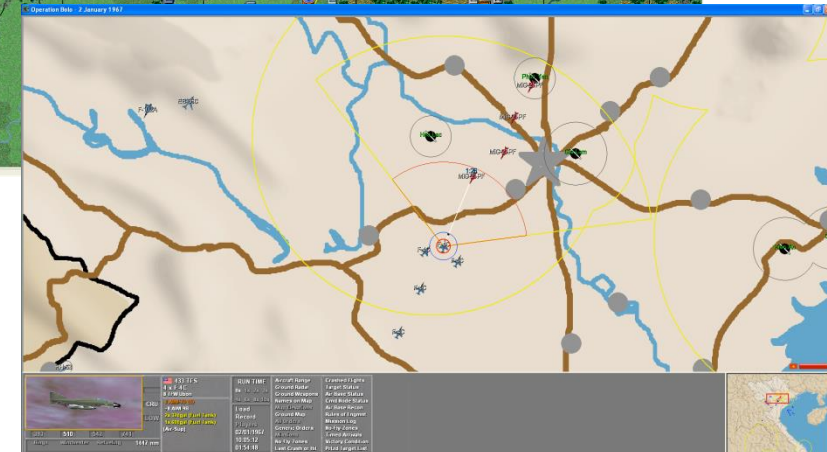
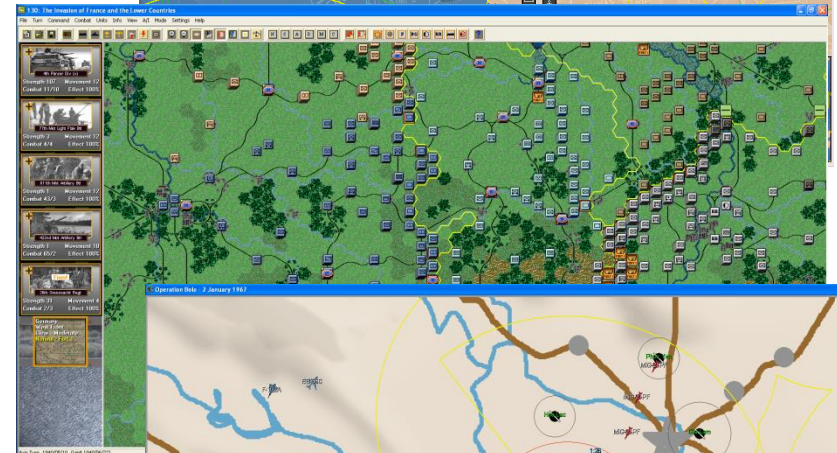


Molecule View of Concept Associations

AI in Cyber Wargaming provides much needed insight into system interactions and human activities

- Simulation and Modeling of Large Systems
 - Training
 - Strategic and Tactical Planning
-
- **Cyber warfare modeling and simulation development for the Air Force Research Laboratory**
 - **War Games deployed for training at the Air Force Squadron Officer College**
 - **Every officer uses the game to study strategic and tactical measures**
 - **Can be used for analyzing what-if scenarios including asymmetric warfare and the use of advanced weapons systems**
 - **War game Artificial Intelligence Development in conjunction with John Tiller Games**
 - **Many commercial titles published**

The First Blitzkrieg won the Wargamer Award for Excellence



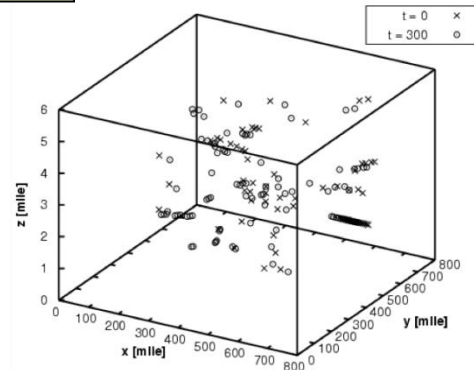
Vantage: Sensor Network Definition, Analysis and Optimization

Simulations provide a means to study new systems and the impact they will have on current and future scenarios

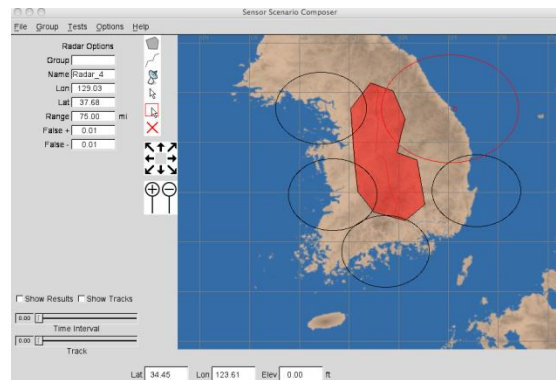
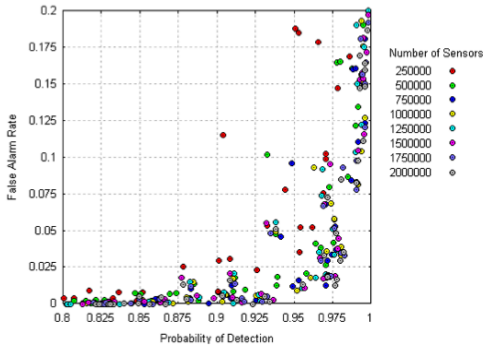


Users define scenarios using Vantage tools

Modeling



Analysis



Visualization of Results

- Model Large-scale Networks of Sensors
- Compose High-density Sensor Networks of Heterogeneous Sensors
- Analyze and/or Determine Sensor Placement and Effectiveness
- Detect and Tracks Targets Over Time (including error rates)
- Evaluate Large Passive Sensor Networks - Defense
 - For detecting hostile aircraft
 - For vectoring interceptors
 - For activating SAM sites
 - For guiding SAMs (initial stage)
- Evaluate Large Passive Sensor Networks – Offense
 - How to best vector aircraft
 - How best to select targets
 - Aircraft detection probabilities and times
- Real Time Integrated Simulation
 - Modern Air Power or other simulation on front end
 - Sensor network on Linux cluster on back end
 - Good for interactive exploration / what if's

Data Albums

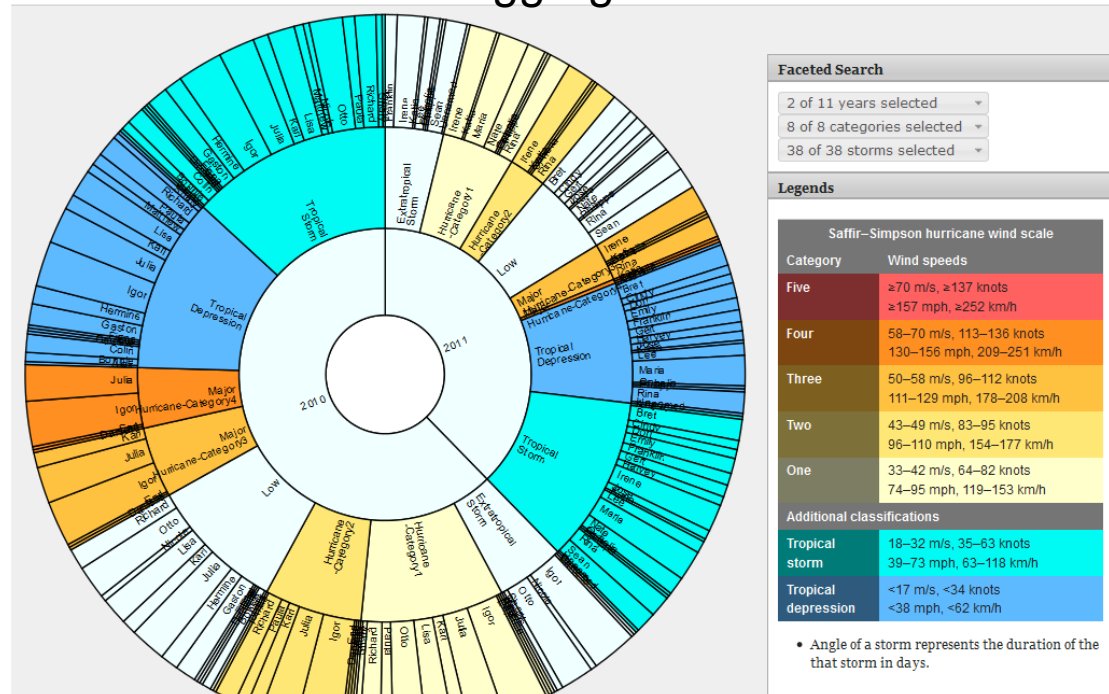
- Supports search, aggregation and access of data and online resources around events
- Automates the gathering of online resources with information filtering
- New ways of displaying search results
 - Infographics – graphic visual representation of information, data or knowledge intended to present complex information quickly and clearly
 - Results enriched with additional information
- Analytics dashboard on the gathered information for events
- Use of semantic technology for relevancy ranking

Compiled collections of information related to a specific topic or event with links to relevant data files. Additional information includes:

Tools and services for visualization and analysis

News reports, images or videos to supplement analysis

Curation provides the author of a Data Album the means to select the aggregated information



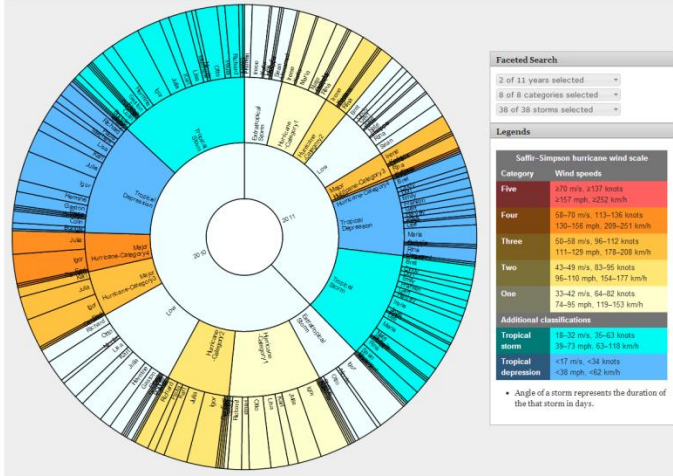
Data Albums

Compiled collections of information related to a specific science topic or event

About the Data Albums

- What is in a Data Album?**
It contains links to relevant data files (granules) from different instruments, tools and services for visualization and analysis, information about the event contained in news reports, images or videos to supplement research analysis, and literature references.
- Value to Researchers?**
To gather relevant data and information for case studies and climatology analysis is tedious and time consuming, current NASA Earth Science data systems are not suited to assemble multi-instrument, multi-mission datasets around specific events. Data Albums addresses this need by automatically gathering all the information and data needed.
- Underlying Technology?**
Data Albums are auto-generated by a reusable and scalable Knowledge Synthesis Engine (KSE). KSE uses not only what information and data resources are relevant but also the semantic relationships between the various fragments of information and data.

Visual faceted search for Hurricanes



Faceted Search

2 of 11 years selected
8 of 8 categories selected
38 of 38 storms selected

Legends

Saffir-Simpson hurricane wind scale

Category	Wind speeds
Five	150-200 mph, 133-177 knots 157-202 mph, 136-174 knots
Four	130-155 mph, 113-136 knots 130-156 mph, 109-121 knots
Three	96-119 mph, 83-103 knots 115-123 mph, 100-108 knots
Two	74-95 mph, 64-82 knots 96-119 mph, 83-96 knots
One	39-55 mph, 34-48 knots 33-42 mph, 29-36 knots

Additional classifications

Classification	Wind speeds
Tropical storm	39-55 mph, 34-48 knots
Tropical depression	<39 mph, <34 knots

• Angle of a storm represents the duration of that storm in days.

Analytics

Box Plot | Interactive | Density Map | Raw Data | Data Summary

Interactivity

YEAR: 2011

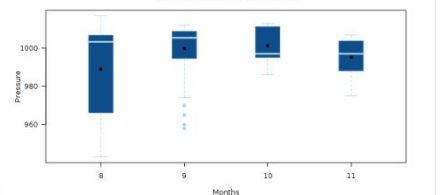
Boxplot Parameters:
Pressure: [Select]

Drill-through-to-detail (Boxplot)
 Drill to Months

Download Box Plot as PNG
Download Plots

Standard Boxplot with Mean

Monthwise Pressure for 2009



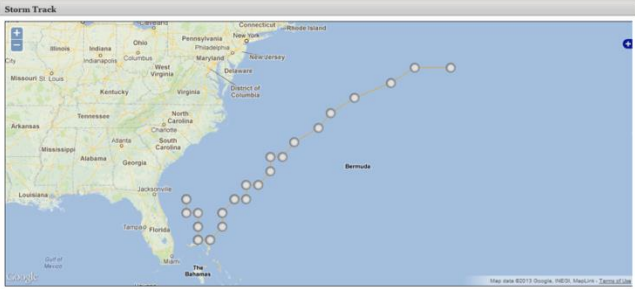
Data Albums

Compiled collections of information related to a specific science topic or event

Box

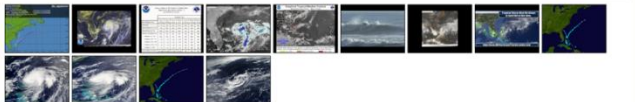
Data Album for Bret (2011)

Storm Track

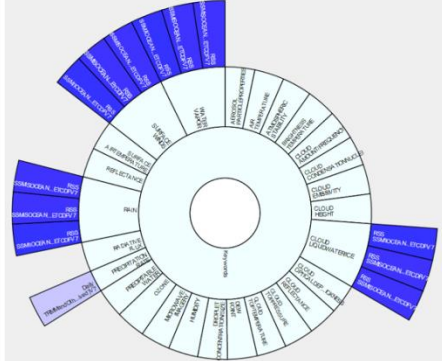


Storm Statistics & Multimedia

Storm Name	Year	Cyclogenesis Date	Cyclops Date	Highest Saffir-Simpson Category	Maximum Wind Speed	Minimum Central Pressure	ACE Index value	Damage	Track Errors	Intensity Errors	Genets
Bret	2011	07/18	07/21	TS	63 mph		2.8456 x 10 ⁴	N/A			N/A



Collections based on Keywords or Instruments



Faceted Search

26 of 26 keywords selected
Select instrument(s)
10 of 10 collections selected

Threshold

Min Rank: 0
Max Rank: 0.19794
Current Threshold: 0.04

Legends

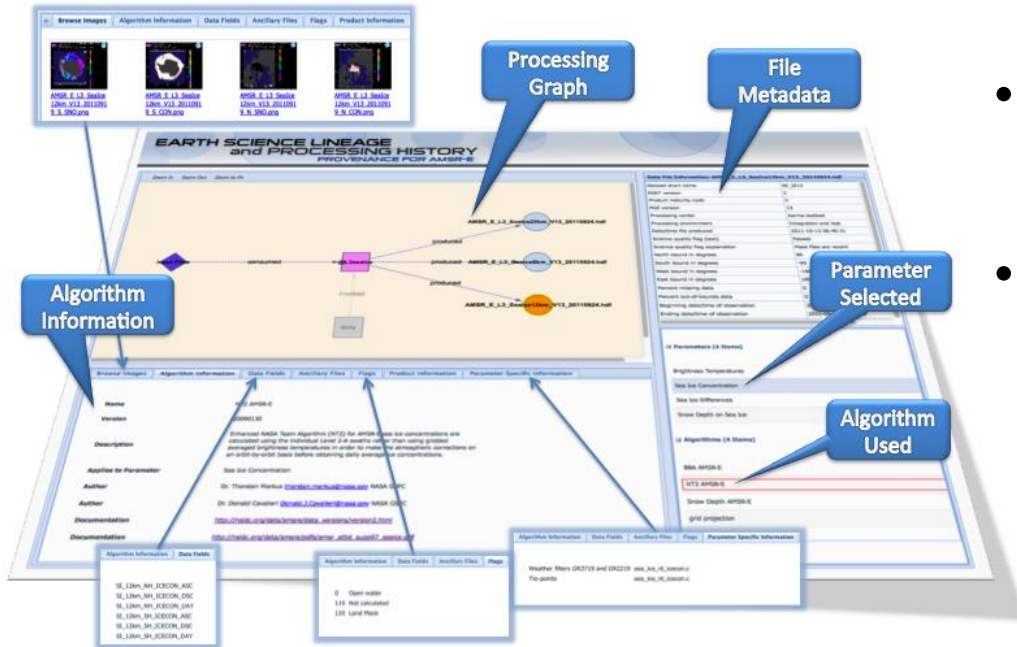
Collection ranking from [Red] to [Blue]
Blue color represents the lowest rank while Red color represents the highest one.

Granules

ID	Name	Date Range	URL
3842_gran_2011.07.17.7.0m	GFSC5A	fp:3842.nascom.nasa.gov/pds/data/4pp/TRMM_33/TRMM_3842_gran_201107170842_gr	fp:3842.nascom.nasa.gov/pds/data/4pp/TRMM_33/TRMM_3842_gran_201107170842_gr
3842_gran_2011.07.18.7.0m	GFSC5A	fp:3842.nascom.nasa.gov/pds/data/4pp/TRMM_33/TRMM_3842_gran_201107180842_gr	fp:3842.nascom.nasa.gov/pds/data/4pp/TRMM_33/TRMM_3842_gran_201107180842_gr
3842_gran_2011.07.19.7.0m	GFSC5A	fp:3842.nascom.nasa.gov/pds/data/4pp/TRMM_33/TRMM_3842_gran_201107190842_gr	fp:3842.nascom.nasa.gov/pds/data/4pp/TRMM_33/TRMM_3842_gran_201107190842_gr
3842_gran_2011.07.20.7.0m	GFSC5A	fp:3842.nascom.nasa.gov/pds/data/4pp/TRMM_33/TRMM_3842_gran_201107200842_gr	fp:3842.nascom.nasa.gov/pds/data/4pp/TRMM_33/TRMM_3842_gran_201107200842_gr

Get Links to All Granules | Get Links to Selected Granules

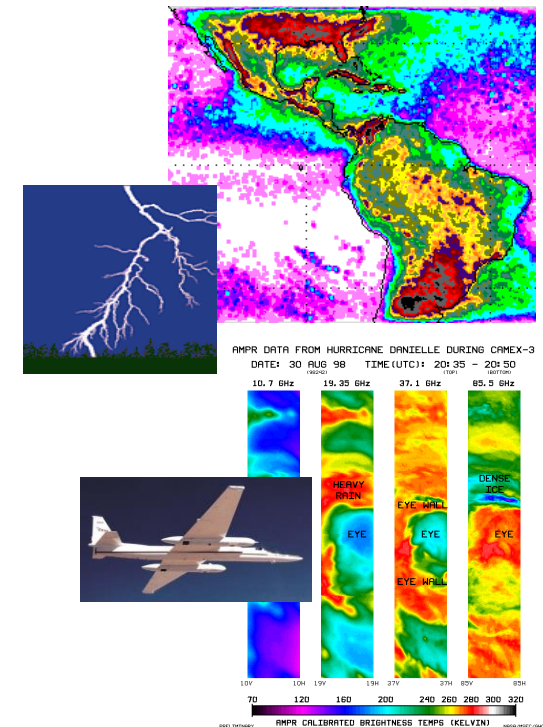
Provenance



- Provenance of hardware, software, and data
- Standards and Interoperability
 - Proposed Open Provenance Model for technology-neutral representation

- Provenance, or “chain of custody”, important to cybersecurity; used to verify trustworthiness and reliability
- Representation, management, presentation; also system engineering and legal, policy and economic issues
- Provenance and context information, such as pedigree and tracking information, may include imagery, data fields, flag values and other information, in addition to a provenance graph and metadata

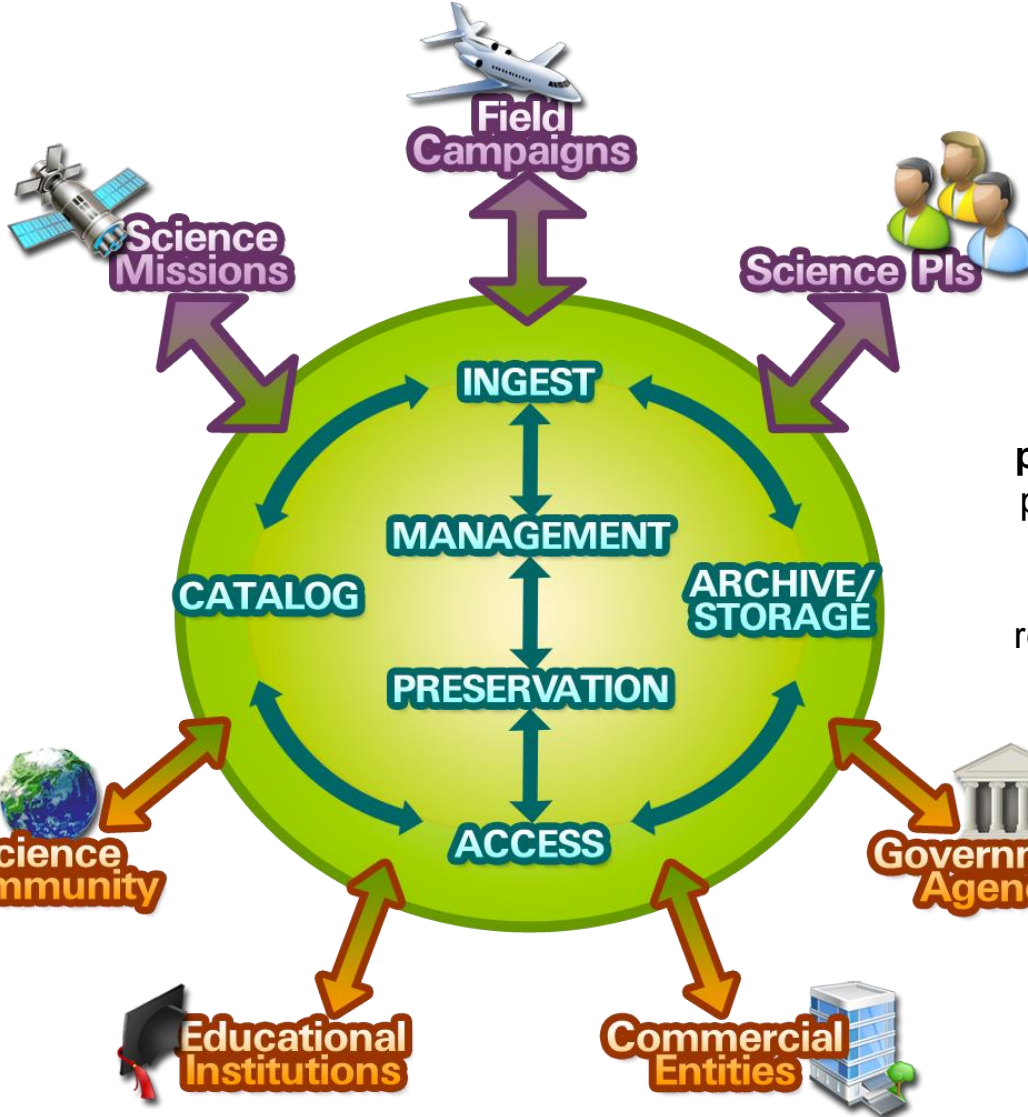
- **Partnership** between NASA and UAHuntsville to apply **advanced information technologies** to a variety of **science data projects**, thereby enabling research and scientific discovery
- **One of twelve full service NASA data centers** providing data ingest, routine and custom processing, archive, distribution, user support, and science data services
 - *Passive Microwave Data*
 - Twenty-five year inventory of data from satellite and aircraft instruments
 - *Lightning Imaging Sensor Science Computing Facility*
 - National lightning data center for the TRMM Lightning Imaging Sensor and validation networks, satellite lightning observations back to 1973
 - *Field Campaigns:*
 - Web-based collaboration for science before, during, and after experiments. Data acquisition, integration, archive and distribution.
 - CAMEX (1998, 2001), ACES (2002), TCSP (2005), NAMMA (2006), TC4 (2007), ARCTAS (2008), GRIP (2010), MC3E (2011), GCPEX (2012), IFloodS (2013), IPHEX (2014), OLYMPEX (2015)
 - *AMSR-E Science Investigator-led Processing System*
 - Generates swath, daily, and monthly products of precipitation, sea ice, water vapor, cloud water, sea surface temps, etc.
 - Near-real-time processing and distribution capability



Data Stewardship

DOCUMENTATION

Capture this information to **create a knowledge base** for our stakeholder communities



CURATION

Work with Science Teams to gather not only data but also **all relevant information**

PRESERVATION

Follow **documented policies** and **engineered procedures** at every step to insure information preservation against all reasonable contingencies

PROCESSING

include **science product generation** and reformatting, algorithm integration and test, interfaces with external providers

INTEROPERABILITY STANDARDS

Ensure that the information is "**independently understandable**" to all stakeholders without requiring experts

* Image from GHRC DAAC

Knowledge Augmentation Services

PROVENANCE

Make the preserved data/information available to all our stakeholder communities with traceability to support authenticity

DATA DISCOVERY

Develop new tools for data discovery, curation and aggregation



FIELD CAMPAIGN INFRASTRUCTURE

Create specialized portals for managing field campaigns and collecting data

DATA USE

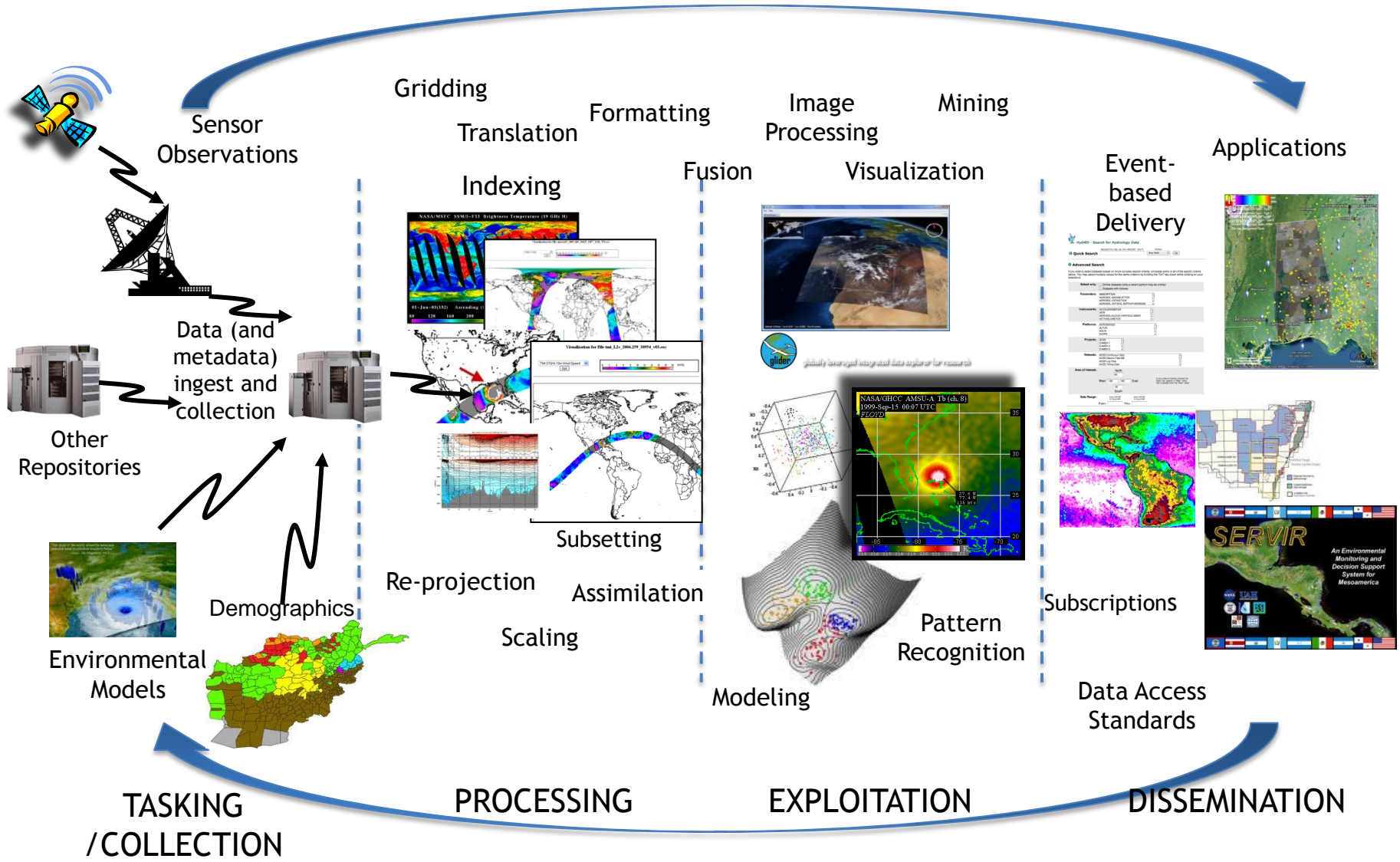
Develop new tools for access, analysis and visualization

INFUSING CUTTING EDGE INFORMATICS

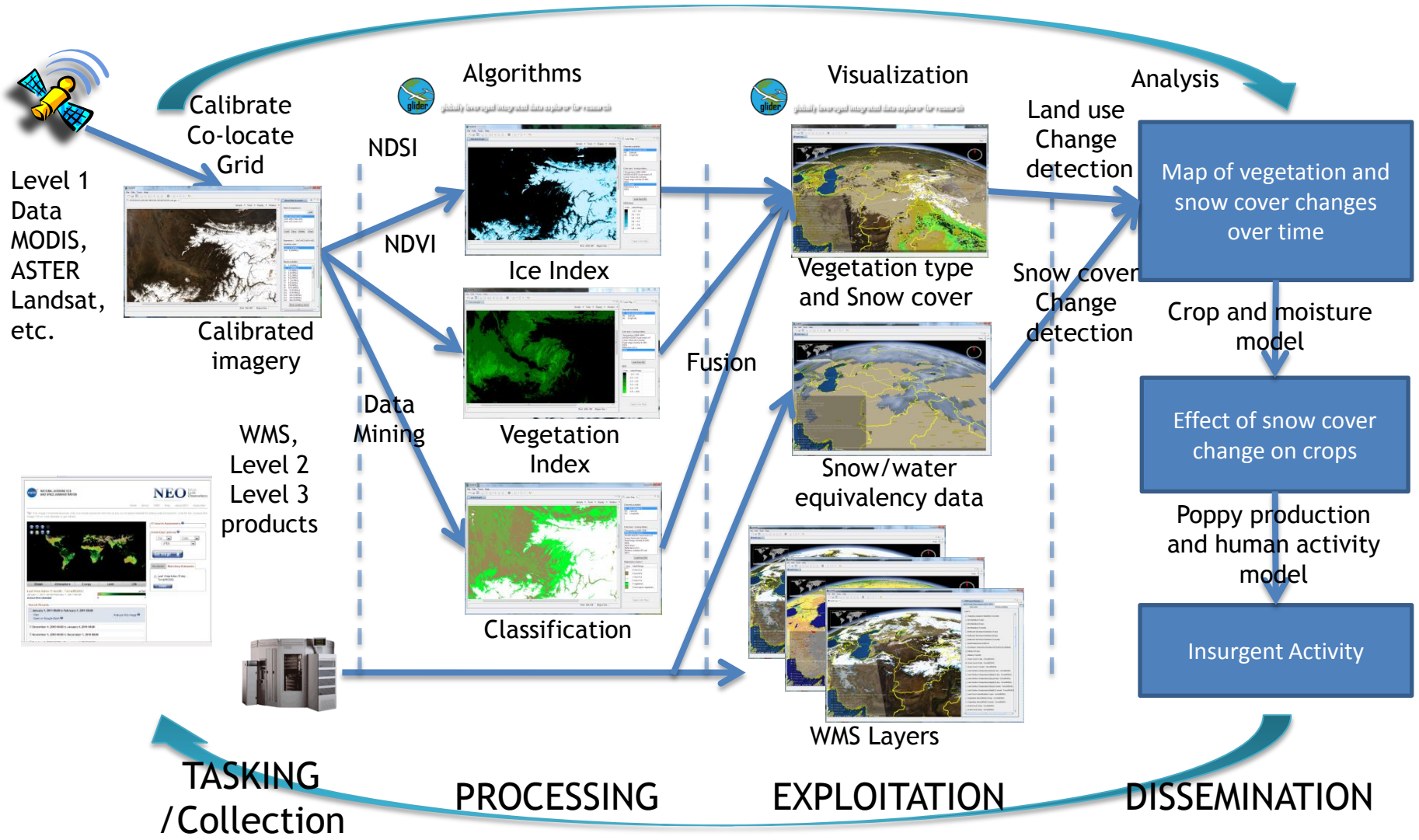
Research new approaches and technologies and infuse them into operational processes

* Image from GHRC DAAC

Conceptual Framework for Multi-source, Multi-function Analysis



Conceptual methodology for examining the linkages between snow cover and insurgent activity in Afghanistan



Cyberinfrastructure Components

**SOFTWARE
TOOLS**

ANALYSIS

MODELING

DATA

COMPUTATIONAL

PEOPLE

**DATA
MANAGEMENT**

NETWORKS