Advancing Collaborative Connections for Earth System Science
Abstracts of selected proposals.
(NNH07ZDA001N-ACCESS)

The ACCESS Cooperative Agreement Notice (CAN) solicits projects that provide strategic, near-term improvements in NASA's Earth science data and information systems by leveraging existing technologies. The specific goals of the 2007 ACCESS announcement is to enhance existing or develop new data system tools and services that improve the access, use, and interoperability of NASA's Earth science data in support of Earth science research and analysis. A special focus on web-based "Service Oriented Architecture" (SOA) tools and services was included.

A total of 30 proposals were received for this announcement. All proposals were peer evaluated using a combination of mail and expert panel review. The Earth Science Division of NASA's Science Mission Directorate selected 10 proposals for two-year awards pending satisfactory budget and work plan negotiations. Some of these awards will be partially funded. These projects will help to further improve NASA's Earth Science Division's heterogeneous and distributed data and information systems.

Philip Callahan/Jet Propulsion Laboratory
Web-based Altimeter Service and Tools

We propose to implement an Altimeter Service (AS) using web-based tools to allow frequent and easy updating and customization of altimeter data products such as the TOPEX GDR with retracking (RGDR). Under our OSTST proposal for TOPEX retracking, we will make a stand-alone version of this Retracked GDR (RGDR) with new precision orbits and Jason-1 geophysical models. However, it is likely that many investigators will have other new orbits and improved geophysical models that could enhance the product or be useful for some specialized research purposes. In order to spare other investigators the difficulty of producing an entire GDR product, we propose to develop web-based tools to allow the creation of links between parts of the RGDR and additional new fields. The links will allow users to dynamically create a residual sea surface height with various orbits, tides, and other geophysical fields produced by cooperating investigators to replace the values provided on the RGDR. This will allow for simple updating or testing and validation of alternative models by many users. To facilitate regional GDR processing and analysis, every dataset will be space/time subsettable and plottable, including the input base GDR, all available component models, and all reprocessed (updated) GDR versions. The proposed development will support Climate Variability and Change science by enhancing studies of global sea level rise and ocean currents. It will also allow local oceanographic studies.

Sara Graves/University of Alabama in Huntsville
Service Mashups for Mining and Analysis with World Wind

NASA's strategic objectives state the need to improve the scientific and technological capabilities of the nation. Today's researchers and students have a difficult time fully exploiting the large volumes of NASA satellite imagery data, in part due to lack of tools that can provide the full spectrum of visually browsing remote data, analyzing a part or all of the imagery and extracting thematic information using advanced mining capabilities. To address this problem, we propose to integrate NASA World Wind, Interactive Visualizer and the Image Classifier for Satellites (IVICS) and the Algorithm Development and Mining (ADaM) toolkit using service mash-ups. By leveraging existing mining and image processing services, the proposed integration will be both effective and efficient.

The goal of this proposal is to integrate World Wind, IVICS and ADaM into a single seamless end user application, modifying the existing components as needed and integrating them via service mash-ups. The tool will be evaluated in research and academic settings using specific science case studies and student projects, focusing on improved visualization and classification of aerosols in combined MODIS/CALIPSO data. The team will refine and customize the tool based on feedback, and make it widely available for diverse applications with a wide variety of NASA imagery.

World Wind, IVICS and ADaM are well established, mature tools with distinct, complementary capabilities. Synergistically integrating these individual tools will result in a complete, comprehensive, easy-to-use end user application for visualization, analysis and thematic information extraction from NASA imagery.

Hook Hua/Jet Propulsion Laboratory
Web Services-enabled Tool for Distributed Custom Level 2 Data Subsetting and Level 3 Data Summarization

We will develop a distributed service-oriented tool that will be used to study long-term and global-scale trends in climate, water and energy cycle, and weather variability. We will provide to the community a Web Services tool that enables scientists to access the NEWS Level 2 data merged from multiple instruments in the NASA's A-Train satellite constellation and to customize the conditional subsetting of the Level 2 data and the production of Level 3 data from pre-summarized Level 3 data according to their specific needs. We will facilitate the transparent access and manipulation of heterogeneous and distributed data by science users via Web Services on a multitude of programming environments. The tool will be comprised of (1) algorithms that subset Level 2 data and generate and transform Level 3 data, (2) Web Services that expose the algorithms and facilitate data access, (3) client-side modules that enable users to discover, access, and manipulate the distributed data, and (4) data crawler and catalog. We will demonstrate the applicability of the proposed tool in studies of climate systems and climate change by capturing long-term global trends and covariabilities of multiple instrument data from A-Train satellite constellation.

This work directly addresses the objectives of the ACCESS solicitation in three ways. First we will improve NASA Earth science data interoperability to facilitate the transparent access and manipulating of heterogeneous and distributed data by providing a Web Services tool that enables science users to access merged data from multi-instuments in the NASA's A-Train satellite constellation. A-Train sensor data provides detailed record of simultaneous observations of temperature, water vapor and cloud properties. Currently there does not exist a capability to discover and access data from multiple instruments in the A-Train as merged multi-parameter data sets. Second, we will increase users' ability to customize their discovery, access, delivery, and manipulation of NASA Earth science data by providing a Web Services tool that allows science users to customize the conditional subsetting of Level 2 data and the production of Level 3 data on a multiple of programming environments. We will create client-side modules that are general enough to encompass various needs of scientists and are easy to interface with the scientific code that science users are using. Third, we will deploy existing Earth science research analysis tools and software using a web-based Service Oriented Architecture paradigm by deploying a well-established statistical summarization tool as a Web Service. The summarization tool produces Level 3 data that preserve the critical instantaneous relationship between parameters and comprehensively characterize the long-term and global-scale trends and covariablity of multiple parameters.

Charles Ichoku/ESSIC/UMD at NASA/GSFC
Integrated Validation, Intercomparison, and Analysis of Aerosol Products from Multiple Satellites.

Among the known atmospheric constituents, aerosols constitute the greatest uncertainty in climate and air-quality research. Although several NASA and other satellite-borne sensors retrieve aerosols, identical aerosol parameters from different sensors often disagree, leaving users confused as to which sensors to trust for answering important science questions about the impact of aerosols. The Aerosol Robotic Network (AERONET)provides the most globally representative ground-based aerosol optical thickness (AOT) and other data. Part of the success of aerosol retrieval from MODIS is due to extensive validation with AERONET using MODIS Atmosphere Products Subset Statistics (MAPSS); a powerful software/database system, which continuously derives collocated statistics (mean, median, standard deviation, etc.) of MODIS (50x50-km) data subsets centered over each AERONET station, and those of AERONET data segments within +30 min of each satellite overpass. MAPSS data are provided online for easy access.
For this proposal, we plan to extend MAPSS to incorporate essential aerosol (AOT, fine mode fraction, single scattering albedo, angstrom exponent, Aerosol Index, etc.) and related (water vapor, cloud fraction, etc.) data from Terra-MODIS, Aqua-MODIS, Terra-MISR, Aura-OMI, POLDER, CALIPSO, and AERONET. The integrated data will be provided online through the NASA-developed Giovanni internet tool (http://giovanni.gsfc.nasa.gov) for quick analysis and download. In addition, combined time series plots of average regional AOT from AERONET and each of the satellite sensors will be provided online in graphic form and updated daily, to enable both the data

producers and users to monitor in near real-time, the agreement between the sensors and AERONET, which is the standard baseline for AOT. Since the data will reflect the unique (spectral, spatial, angular, polarization, or profiling) capabilities of the sensors, we will conduct at least two main studies in selected regions with different aerosol types: (1) evaluate the products and provide guidance as to which is more reliable or at least an indication of why the discrepancies exist, (2) investigate the variability of important aerosol properties that are best studied using multi-sensor data.
This proposal responds directly to at least one of the important science questions ("How is atmospheric composition changing?") in the NASA Science Strategy. It also addresses the core of the EOS program element of this solicitation, which emphasizes the "provision of well-calibrated, multiyear and multisatellite data and product series", and specifies that "studies that utilize multiple sensors and platforms are of particular interest".

Steven Kempler/NASA/GSFC
Utilizing 3 Dimensional Data Views to Access Data and Discover Relationships Between Multiple Heterogeneous Data Sets Along the A-Train Tracks

NASAs A-Train, comprised of a succession of US and international satellites that follow each other, seconds to minutes apart, across the local afternoon equator, provides great opportunities to increase the number of observations, validate observations, and enable coordination between science observations, resulting in a more complete "virtual science platform". (Kelly, 2003). The A-Train consists of the following satellites, in order of equator crossing: OCO, EOS Aqua, CloudSat, CALIPSO, PARASOL, and EOS Aura all in the same sun-synchronous orbit. Unfortunately, when each project was conceived and implemented, formation flying as described above, and the wealth of science that would result, was not considered when data structures, data formats, fields of view, instrument resolutions, etc, were designed for each instrument. Projects like the A-Train Data Depot (ATDD) (Kempler, 2006), and efforts at Colorado State University (Reinke, 2006) and University of Utah (Mace, 2007), have made great strides in providing researchers means to visually explore the A-Trains vertical 'curtains', provide data access, and distribute various datasets along the A-Train tracks, thus bearing the fruits of science promised by formation flying. Not only have these efforts made otherwise heterogeneous A-train data available from one virtual data portal, they have also provided data pre-subsetted and pre-processed to common grids, thus further facilitating the use of the data while removing the burden of this preprocessing from thousands of users.

In addition, A-Train science provides a very unique challenge in that it involves datasets studied specifically in three dimensions. Whereas, Cloudsat and CALIPSO measure vertical profiles, most other instruments measure in the horizontal plane. Some instruments, like MLS and AIRS provide data in 3 dimensions.

The purpose of this project is to develop the tool, A-Train Data in 3 Dimensions (ATD3D), that employs the latest 3 dimensional visualization technology to explore and provide direct data access to heterogeneous A-Train datasets, “operationally”, along, and on either side of the A-Train tracks, that emphasize the multi-dimensional significance of

cross instrument A-Train data. Google Earth provides the foundation for displaying vertically and horizontally oriented datasets. For example, visualizations such as Cloudsat and CALIPSO vertical curtains, MODIS cloud top pressures, and OMI cloud pressure, can reveal cross instrument data signatures not otherwise easily detected. In addition, ground meteorology, such as surface temperature, pressure, winds, and rainfall, will be available to inter-compare atmospheric measurements with synoptic weather conditions. Utilizing Google Earth's ability zoom, pan, tilt, and rotate, provides users to the opportunity to examine features prior to downloading data, perhaps unnecessarily. Once data of interest is discovered, users will be able to access the specific datasets from the archive in which the raw data resides through web services.

Tying A-Train data on Google Earth with the ATDD increases users ability to discover, access, manipulate and analyze A-Train atmospheric data. This project will make use of: Google Earth for information discovery; The GSFC Earth Sciences (GES) Data and Information Services Center (DISC) Giovanni front-end for data search; The GES DISC S4PA for serving local data, and; OpenDAP for access to remote data. All services are TRL 9. This project will facilitate the further utilization of NASA Earth science through the integration of publicly available tools.

Ramakrishna Nemani/NASA Ames Research Center
TOPS: Extending access to NASA data and model results for ecosystem studies.

The Terrestrial Observation and Prediction System (TOPS)[http://ecocast.arc.nasa.gov] is a flexible modeling system that integrates ecosystem models with satellite and surface weather observations to produce ecosystem nowcasts and forecasts useful in natural resources management, public health, and disaster management. Over last three years, TOPS has served 1.5TB of custom-processed data to users from the federal and state governments, university laboratories and the private sector. However, current TOPS user interface is purely static and doesn't allow for user interaction. Users are only served through a standard ftp repository, and without the ability to query or preprocess the data. Any custom processing requests are handled on a case-by-case basis by TOPS team members and it involves some level of manual system interaction. This approach does not scale well and we will not able to support more users despite the increasing interest in TOPS data for research and applications. We propose to augment the current TOPS functionality by providing extensible, flexible, transparent access and discovery of data, and on-line on-demand analysis services through integration of existing TOPS utilities and databases together with existing open source tools, protocols, and standards for interoperability. The users will be able to perform custom visual on-line data analysis and manipulations by chaining together services and datasets provided by TOPS using drag-and-drop interface in an interactive and intuitive way. The access to the data and services will be available both through a web-based user interface based on Ajax and DHTML, as well as through published interfaces to the underlying Service Oriented Architecture (SOA) so that users can build their own interactive or batch applications based on direct machine-to-machine interfaces. The proposed enhancements to TOPS will enhance access to NASA data, model results and services to our current users as well as a growing community conducting climate change impact and ecosystem studies.

Rahul Ramachandran/University of Alabama in Huntsville
SAM: Smart Assistant for Earth Science Data Mining

Scientific data mining is a very powerful means for automated knowledge extraction from the ever-increasing volumes of science observations and model output data available. NASA's Second Data Mining Workshop found that maturing data mining techniques show "potential for significantly expanding the scientific understanding of NASA's Earth science data." However, this type of tool has generally been difficult for domain scientists and students to fully exploit without extended learning curves. And even data mining specialists may not be familiar with the full range of components in a mining toolkit, so potentially useful mining strategies may be ignored. To facilitate exploitation of these promising techniques by the increasingly IT-sophisticated NASA Earth science community, the University of Alabama in Huntsville leads a collaborative team in proposing to leverage Semantic Web technologies to build a Smart Assistant for Mining (SAM) and to deploy it for use at two data centers. This project will reuse an existing toolkit of data mining web services designed specifically for the analysis of NASA data in a web-based, service-oriented architecture. It will also leverage and extend an initial ontology describing data mining services, with links to other ontologies describing the Earth science problem domain and relevant data sets. The new SAM user interface tool, which integrates semantic reasoning into a traditional workflow composer, will allow users to discover available data and services, assist users in composing mining workflows, and invoke them to perform the desired analysis. SAM will provide a useful tool to assist researchers in creating data analysis and mining workflows for targeted Earth science problems in the Climate Variability and Change Science Focus Area. It will also position these services for integration with many other science data services in the Semantic Web Services context, pointing the way toward increased science return from NASA data.

Robert Raskin/Jet Propulsion Lab
Virtual Oceanographic Data Center (VODC)

On-line oceanographic mission data sets are published by numerous state, federal and international data centers typically with disparate software interfaces and information models by which to query and retrieve information. The process of coordination, search, and discovery of these data sets has become increasingly burdensome and it is time consuming and costly to locate and extract data and metadata from unfamiliar data sources. The development of a unified, systematic approach for searching, indexing and discovering oceanographic data sets and their respective metadata across organizations can do much to alleviate the time and complexity costs borne by data users. We will develop a Virtual Oceanographic Data Center (VODC) as a "one-stop" Internet presence for searching, extracting, and fusing data from oceanographic data sources. We will apply the Object Oriented Data Technology (OODT) middleware (2003 Runner-up NASA Software of the Year) to develop a standards-based integrated catalog of oceanographic data.

Julienne Stroeve/NSIDC/University of Colorado
Cryospheric Change Analysis Web Services: Application of a Suite of Tools to Greenland Ice Sheet Processes

NSIDC and a group of collaborators propose to create advanced multi-faceted web services for data access, browse, overlay, manipulation, and delivery of data sets important for cryospheric research. Our focus area for this work is the Greenland ice sheet. However, the approach is extendable to any region of interest.

The proposed work leverages many of our existing subsetting, gridding, projection, and visualization tools into modular services, invoked through a web-hosted geospatial data management system, or as individual services. Our main target data sets are: MODIS and AVHRR, AMSR-E and SSM/I, GLAS and ATM, CERES, and QuikSCAT, already generated and available at NSIDC or from other sources on the web.

We address several widely perceived needs in cryospheric and earth science research:
o Specific community need - Understanding the processes controlling the mass balance of Greenland's ice sheet and its contribution to global sea level through analysis of cryospheric variables such as surface temperature, albedo, melt extent, and surface elevation. Such analyses are critical to ongoing programs such as the NASA funded Program for Regional Climate Assessment (PARCA) and the NSF funded Arctic Observing Network.
o Climate science need - Climate change studies require the ability to integrate and compare multiple data sets, and to easily observe parameter trends in time-series analysis. We propose a new approach that allows for rapid time-series browse and analysis.
o General earth science need - Addresses the increasing challenge of efficient access to and analysis of large and widely-distributed remote sensing data sets for individual scientists.

The services we propose directly address significant data access issues that currently impede progress in our understanding of the climate change response of Greenland's ice sheet. By providing a simple and intuitive mechanism for browsing, analysis, and delivery of diverse data sets, we enhance the value of these data. By leveraging existing in-house tools and data sets, and designing tools to use remotely available data storage and documentation, we simplify the overall task and make maximize use of existing infrastructure.

Thomas Yunck/Jet Propulsion Laboratory
A Climate Virtual Observatory (CVO): Online Data Fusion and Analysis for Climate Variability and Change

With this project we propose to: (1) assemble the components of the GENESIS/SciFlo Earth science analysis system developed under NASA's REASoN program into an online, interactive Climate Virtual Observatory; and (2) employ CVO in a critical recalibration

of historical remote sensing data to reveal recent temperature trends throughout the atmosphere.

CVO will specialize in atmospheric data from a variety of spaceborne sensors, including AIRS, AMSU, MODIS, MISR, and AMSR-E on NASA's Terra and Aqua satellites, and GPS radio occultation (GPSRO) data from CHAMP, SAC-C and COSMIC. Several other data sources will also be accessible. CVO specifically targets multi-sensor studies in regional and global climate variability and change. The datasets and analysis tools it offers will also be of use in studies of atmospheric composition, weather, and carbon and water cycles. CVO will be accessible online through its own URL, as well as through the ESIP Federation's Earth Information Exchange.

Our science team will apply CVO in a fundamental recalibration of nearly a decade of atmospheric temperature data from AIRS, AMSU and MODIS on Terra and Aqua, and earlier AMSU data from two NOAA satellites. This will be achieved using the precise and absolute data from NASA's GPS occultation instruments on CHAMP, SAC-C and COSMIC. The recalibrated radiometer data will reveal with unrivalled clarity the recent temperature trends at different altitudes and may resolve some fundamental questions as to how the atmosphere as a whole is responding to climate change. The corrections will be applied both in the present and retroactively, even before GPSRO data were available. From this work we will publish a comprehensive survey of corrected atmospheric temperature trends from ~5 to 30 km altitude for the full globe, five broad latitude bands, all continents, and selected sub-continental regions.