NASA Science Mission Directorate
Research Opportunities in Space and Earth Sciences
Advancing Collaborative Connections for Earth System Science
Abstracts of selected proposals
(NNH11ZDA001N-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 objectives of the 2011 ACCESS announcement are (1) improvements in user's ability to efficiently discover, find, access, and readily utilize useful science content from NASA's increasingly large volumes of multi-mission, multi-instrument Earth science data; (2) tools that improve and expand the accessibility and usability of NASA's Earth science observational data for the modeling and model analysis communities; and (3) other tools and technologies that enhance the accessibility and usability of Earth science data and extend the reach of NASA's Earth Science Division IT investments to new users and communities.

A total of 37 proposals were received for this announcement. All proposals were peer evaluated using an expert panel review. The Earth Science Division of NASA's Science Mission Directorate selected 12 proposals for two-year awards with a total funding of approximately $7.8 M. These projects will help to further improve NASA's Earth Science Division's heterogeneous and distributed data and information systems.

Bruce Caron/New Media Studio
NASA Science on Drupal Central

NASA Science on Drupal Central

Drupal hosting, administration, UI work, and service- and tool integration on Drupal sites are already widespread activities for NASA Earth science. Many NASA Centers, funded projects, and other partners (including university labs and departments) are in the process of, or have completed the migration of their content onto Drupal.

Drupal has become a content management system of choice for several reasons. It is an open-source platform with many thousands of contributors to its core code and extensions, and it has been around long enough to address many issues of security and robustness of value to NASA.

The proposed NASA Science on Drupal Central (NSODC) project would offer key support for, and a centralized knowledge base about the effective use of Drupal. Science on Drupal will deliver a suite of online environments (social networks, code versioning, collaboration environments, etc.) where NASA scientists and technicians can share their Drupal lessons learned, register their code contributions, discuss issues of NASA-specific common interest (e.g., migrating SOA clients, porting workflow engines and data provenance capabilities to Drupal platforms, integration with IDL and Matlab, etc.), and search, find, and reuse NASA-funded Drupal code (on the Science on Drupal Github

repository or elsewhere). While the NSODC will be focused firmly on science and data tools on Drupalwill, the project will also provide a home for more general, Drupal-wide knowledge sharing in support of Drupal site administrators across NASA Earth science.

The NSODC will also build and share tool frameworks as Drupal modules. These can be reused and customized by others to accelerate NASA Earth data tool development on Drupal. Examples include taxonomy, metadata authoring, cloud service integration, and catalog searching. other examples will include extending the existing the Talkoot analysis suite to handle different workflow engines and clients and calling IDL on the server-side to display an image or graph. RDF is now a part of Drupal 7 core, so support for NASA semantic web efforts on Drupal will also be provided.

Beyond this new code, the NSODC team will help organize and lead Drupal camps and code sprints at ESIP Summer Meetings. The team will engage the Drupal community of purpose in active online discussions, supplemented by Drupal events at ESDSWG, ESIP, AGU, and other meetings. These activities will be guided by the code- and knowledge needs expressed by NASA Drupal users.

Tools and code are one half of the effort. The other half is community building and support. NSODC will engage NASA Drupal code developers and site administrators as a community of purpose, and provide avenues of communication, collaboration, and resource sharing. The (Drupal-based) DigitalOcean science social and media networking platform will provide rapid communication, group resource sharing, and reputation tools. The collaborations will be facilitated on the ESIP Federation Open Atrium platform. The resulting code will be hosted on Github, licensed for NASA reuse.

In combination, the code and community deliverables "push the needle" for NASA Drupal use. Significant time and cost savings will result. Now is the right time to bring together the hundreds of independent Drupal coders, themers, and administrators who are already (or planning to be) tasked by NASA to implement Drupal. The proposed effort provides all of the necessary capabilities to take NASA Drupal to a new stage driven by collective intelligence.

The NASA Science on Drupal Center project will be led by Bruce Caron at the New Media Research Institute and Rahul Ramachandran at UAH, and will include Martin Landsfeld at NMRI.

Chris Currey/Langley Research Center
Multi-Instrument Inter-Calibration (MIIC) Framework

Improving accuracy of the satellite-based Earth Observing System is a crucial objective for climate science. Climate quality measurements require sophisticated satellite instrumentation to monitor the onboard calibration, currently available on very few sensors. Inter-calibration is the process of comparing co-located measurements with matched viewing geometries from different sensors on separate spacecraft to improve target instrument calibration and remove temporal calibration trends. Inter-calibration ties target instrument calibrations on one spacecraft to highly accurate, preferably SI-

traceable, reference instruments on another spacecraft. Low Earth Orbit (LEO) and Geostationary Earth Orbit (GEO) instruments with similar spectral responses are compared by analyzing co-located pixels using a line-of-sight matching technique near orbit intersections. Error analyses (not part of this proposal) specify achievable inter-calibration accuracies for various instruments and sampling criteria in time/space/angle dimensions.

Two international organizations are trying to establish best inter-calibration practices for GEO and LEO sensors. The Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV) promotes high quality calibration and validation of infrared and visible optical data. The World Meteorological Organization (WMO) Global Space based Inter-Calibration System (GSICS) community analyzes methods and algorithms to improve the calibration of operational instruments in support of climate monitoring. The development of a common framework to improve access to co-located data with matching viewing geometries from multiple instruments will enable independent calibration groups to advance the inter-calibration of current sensors such as GEO, AVHRR, MODIS, and VIIRS. Furthermore, this framework will be designed to be readily adapted to handle the increased data volume from future passive remote sensing Decadal Survey and international missions, including hyperspectral reference instruments such as Climate Absolute Radiance and Refractivity Observatory (CLARREO) or Traceable Radiometry Underpinning Terrestrial and Helio studies (TRUTHS).

The Multi-Instrument Inter-Calibration (MIIC) Framework provides a real world solution to instrument teams responsible for calibration and validation of target instrument data. The MIIC Framework allows efficient access to reference data from existing Distributed Active Archive Center (DAAC) servers. Inter-calibration (IC) events from multiple spacecraft for a given time period (typically 1 month) are automatically predicted from the specified sampling criteria and orbit crossings. The event specifications are inserted into an Extensible Markup Language (XML) inter-calibration plan. For each event inter-calibration algorithms are executed on remote servers using Open-source Project for Network Access Protocol (OPeNDAP) server-side functions prior to delivery of the data to the instrument teams for further analysis. Algorithms such as convolution over sensor point spread functions and spectral response functions are run on remote data servers. This capability saves months of time downloading extraneous data and reduces local processing and disk storage resources typically required by instrument teams to perform these functions. In addition, this framework will help science teams distinguish trends from calibration artifacts in retrieved parameters such as cloud properties. The goal of this effort is to re-use existing science algorithms and information technology software to increase the interconnectedness of existing distributed information systems and the quality of datasets.

William Emanuel/Pacific Northwest National Laboratory

Enabling Centralized Access to Land Cover Data for Climate Change Integrated Assessment Modeling

The proposed activity will develop and demonstrate data systems that enable effective and efficient use of NASA satellite remote sensing data to derive global land cover data

required by climate change integrated assessment (IA) models. Currently, integrated assessment models rely heavily on land cover data derived from Advanced Very High Resolution Radiometer (AVHRR) measurements collected during the 1992 -1993 period in order to characterize crop and pasture cover as well as distributions of naturally occurring vegetation. The project will assemble and make available through centralized access NASA remote sensing data needed to improve land cover characterizations within integrated assessment models and distribute new data as they are developed by the IA modeling community.

Climate change integrated assessment models relate human factors and activities, such as demography, energy use, technology, the economy, agriculture, forestry and land use to greenhouse gas emissions, other perturbations to the climate system, and to the resulting radiative forcing of climate change. Such models are used in diverse studies to investigate potential pathways of future climate change, vulnerability, and options for mitigation or adaptation. Among the most prominent products derived using integrated assessment models are emissions scenarios or other representations of potential future pathways of human activities as they affect the climate system. Such scenarios are the basis for major climate change assessments such as those of the Intergovernmental Panel on Climate Change (IPCC).

The project will make remote sensing data available to IA modeling groups in global mosaics and in formats that are compatible with other data used with IA models. Efficient access to NASA remote sensing data will enable modeling groups to derive improved land cover data, including continuous fields for crop, pasture, urban and irrigated land, as required by IA models. Remote sensing data in these formats and with global coverage as well as prototype land cover products produced by this project in cooperation with the IA modeling community will be applicable to global modeling and analysis of terrestrial ecosystems for biogeochemistry, climate change impacts, biodiversity, and other similar studies at larger spatial scales.

Consistent use of remote sensing data by different groups collaborating in the development of land cover data products for use with IA models is important. Even in cases where different investigators or groups derive alternative IA model data sets, the data are far more useful if the underlying remote sensing data are consistent, allowing isolation of differences in assumptions and processing that are most relevant to IA models. Recognizing this use of IA models, the project will also develop and demonstrate a data distribution system especially well-suited to the IA modeling community, which requires version control to insure consistency between analyses by different modeling groups worldwide. Integrated assessment modeling groups and other interested users will also be able to acquire data served by a version control server.

Expected products of the proposed project include global remote sensing data for use in developing land cover characterizations within IA models, prototype land cover data sets developed by the project and data sets developed by the IA community, centralized access to land cover data by the IA community and other users, and a system for data distribution with version control and data file synchronization between servers and user

computers. Results of the project will be reported at major meetings, described technical reports, and summarized in journal publications.

Hook Hua/Jet Propulsion Laboratory
Collaborative Climate Model and Observational Data Services (CCMODS)

Retrospective-analyses (reanalyses) and free-running climate models are critical tools in studying weather and climate variability and change. An important--but still incomplete--requirement is long-term continuous comparisons between models and observational data. From our recent peer-reviewed NASA-funded NEWS and MEaSUREs projects, we have amassed a large collection of multi-sensor A-Train data and analysis capability spanning almost a decade. Level 2 comparisons of these data with models are an untapped science opportunity. For example, the MERRA project has generated a long-term record of the hydrological cycle, and fits well with model comparisons using our NEWS and MEaSUREs observations of the atmospheric water cycle. Modelers need data sets (such as NASA's multi-sensor A-Train data) to identify and correct biases in modeled physical processes occurring at many spatial or temporal scales. Collocation of satellite-measured water vapor, clouds, and temperature onto model grids is necessary for evaluation of model simulations of physical processes. In addition, scientific advances can be made by long-term global monitoring of climate-relevant processes quantified by the multi-year A-train record. The major difficulty with quantifying atmospheric phenomena over the lifetime of the A-train/NASA EOS era is in efficient sub-sampling, quality control, and automation to obtain the necessary data (both from NASA data and modeling centers). Scientists need good dataset preparation and screening to properly compare Level 2 observations to sub-daily model grids. Tools and services are needed to streamline scientists' custom processing and analyses to assess and improve models. These are key uses of NASA data sets for improving the understanding of weather and climate variability and change.

We will reuse and integrate existing technologies and data sets (many developed from our prior ACCESS, MEaSUREs, NEWS, and ARRA grants) to reduce interoperability barriers and the unnecessary complexities in accessing, merging, and comparing multi-sensor satellite observations from the A-Train with climate models and reanalyses. Scientists performing model and observation comparisons for climate analyses face hurdles in data discovery, access, subsetting, subsampling, quality screening, regridding, collocation, analysis, and data sharing & collaboration. Our system, Collaborative Climate Model and Observational Data Services (CCMODS), will address these data access and interoperability issues that often exist in NASA Earth Science's distributed and heterogeneous data and information systems. CCMODS will also simplify the voluminous data transfer issue by automating observation and model data assembly, merging, and analysis on the server side. More importantly, these new capabilities address specific science gaps in the model evaluation process. Our science investigators will use this data system to move comparisons of satellite observations to models beyond monthly mean comparisons and simple statistics (mean differences and standard deviations). Our data system will be used to compare instantaneous Level-2 satellite observations of temperature, water vapor, etc. to three-hourly MERRA and ECMWF model grids. Our science team will also use the system to perform detailed model

diagnosis (differences from Level-2 observations and trends) stratified by observed atmospheric processes (cloud scene and precipitation intensity), retaining the full Probability Density Functions (PDFs) of all relevant observed quantities.

Meemong Lee/Jet Propulsion Laboratory
Multi Mission Observation Operator (M2O2)

The goal of the Multi-Mission Observation Operator (M2O2) is to create a streamlined interface mechanism between the atmospheric chemistry model developers and the atmospheric sounding mission data providers by infusing mission-generic observation integration technologies developed under the Advanced Information System Technology (AIST) program. The M2O2 will address a major challenge in utilizing the space-based observations within the atmospheric chemistry modeling and assimilation community, which involves linking between the model analysis and the observed atmospheric state in the level 2 mission data products (L2 data). The state-of-the-practice is to develop an observation operator for each atmospheric component of an atmospheric sounding mission, which often involves laborious data preparation. A wide range of observation operators with their own ad-hoc ways of handling L2 data greatly hinders integration of observations from multiple missions. Developing a generic observation operator that provides the link between the model analysis and mission observations requires an automated "assimilation-purpose" data preparation service and representation coordinate transformation.

The M2O2 will develop a model transformer that shields the cost function analysis process from the mission-specific L2 data and demonstrates the mission-generic global assimilation process for three types of space-based observations (ozone [O3], carbon monoxide [CO], carbon dioxide [CO2], column carbon dioxide[x CO2]). The global assimilation process will be delivered to the GEOS-Chem-Adjoint working group that maintains data assimilation services for atmospheric chemistry community. The M2O2 will also develop a data transformer to provide the model transformer "assimilation-purpose" L2 data, referred to as L2# data, by applying a set of information filters and quality-control filters. The data transformer will be integrated within Goddard Earth System Data and Information Service Center (GES DISC) as a web-service for providing "on-demand" observation information from four types of atmospheric sounding missions, Microwave Limb Sounder (MLS), Tropospheric Emission Spectrometer (TES), Atmospheric Infra-Red Sounder (AIRS), and Atmospheric Carbon Observatory System (ACOS). The model transformer and the data transformer collectively perform as a mission-generic observation operator in the assimilation process, allowing the atmospheric chemistry modeling community to effectively utilize L2 data products from multiple missions.

Our proposal directly addresses the needs identified by the Advancing Collaborative Connections for Earth Science Systems-NASA Research Announcement (ACCESS-NRA) in following three areas:

1) Work with modeling and model analysis communities to effectively find, understand, and appropriately use Earth science observational data. The M2O2 will allow the atmospheric chemistry modeling community to access space-based

observations of O3, CO, CO2, and x CO2 as "assimilation-purpose" data products from GES DISC.

2) Automate the discovery of heterogeneous data meeting customizable criteria based on data content, data quality, metadata, and production. The M2O2-service will automatically apply assimilation-purpose data-quality control for three types of atmospheric constituents from four types of atmospheric sounding missions based on the observation conditions employing the meta data (e.g., cloud, aerosol optical depth).

3) Improve users' ability to mine useful information from distributed, large volumes of heterogeneous data. The M2O2 service will provide a uniform interface to heterogeneous L2 data products from four distributed active archive centers (DAACs), reducing the transferred data volume and increasing the information content for global data assimilation of the respective atmospheric constituents.

Charles Meertens/UNAVCO
Seamless Synthetic Aperture Radar (SAR) Archive for Interferometry Analysis

UNAVCO/WInSAR, the Alaska Satellite Facility (ASF), and JPL propose to collaborate in an information technology and data management development project to design and implement a seamless distributed access system for Synthetic Aperture Radar (SAR) data and derived data products (InSAR interferograms). A seamless SAR archive will increase the accessibility and the utility of SAR science data to solid earth and cryospheric science researchers. Specifically, the project will provide simple web services tools to more seamlessly and effectively exchange and share SAR metadata, data and archived and on-demand derived products between the distributed archives, individual users, and key information technology development systems such as the NASA/JPL QuakeSim and ARIA projects that provide higher level resources for geodetic data processing, data assimilation and modeling, and integrative analysis for scientific research and hazards applications. The proposed seamless SAR archive will significantly enhance mature IT capabilities at ASFs NASA-supported DAAC and UNAVCOs WInSAR and Supersites archives that are supported by NASA, NSF, and the USGS in close collaboration with ESA/ESRIN. As part of the proposed effort, data/product standard formats and new QC/QA definitions will be developed and implemented to streamline data usage and enable advanced query capability. The seamless SAR archive will provide users with simple browser and web service API access tools to view and retrieve SAR data from multiple archives, to place their tasking requests, to order data, and to report results back to data providers; to make a larger pool of data available scientific data users; and to encourage broader national and international use of SAR data. The new ACCESS-developed tools will help overcome current obstacles including heterogeneous archive access protocols and data/product formats, data provider access policy constraints, and an increasingly broad and diverse selection of SAR data that now includes ESA/ERS/ENVISAT (and upcoming Sentinel mission), CSA/Radarsat, JAXA/ ALOS-PALSAR, and DLR/TerraSAR-X satellite data and NASA/UAVSAR aircraft SAR data. The list will continue to expand with NASA/DESDynI further increasing the need to efficiently discover access, retrieve, distribute, and process huge quantities of new and diverse data.

Ramakrishna Nemani/Ames Research Center
NASA Earth Exchange: Improving Access to Large-Scale Data and Computational Infrastructure

The overall goal of the proposed project is to enhance access, discovery and integration of data, models and services for the communities planning on using the NASA Earth Exchange (NEX, Nemani et al., 2011, nex.arc.nasa.gov) for the National Climate Assessment (NCA, http://www.globalchange.gov/what-we-do/assessment). NEX is a new collaborative platform that brings together state-of-the-art computing facility with large volumes (hundreds of terabytes) of NASA satellite and climate data as well as number of ecosystem and climate models. NEX facilitates end-to-end execution of Earth science research projects complete with data acquisition, analysis, model executions and result sharing. The proposed project will be executed in several phases, first we will inventory NEX data holdings and databases developed across number of different projects. We will then provide a unified spatio-temporal schema that will enable us to provide NEX users better access to large number of datasets from many NASA instruments and models. We will also extend our existing OWL-based ontology to encompass quality flags semantics from number of NASA's products to improve process automation and interoperability. Secondly, we will inventory current NEX models, utilities and tools and package them together and provide it as a catalog of services to NEX researchers. Third, we will provide search and execute interface on top of the data and services catalog that will enable users to execute data queries, locate tools, models or libraries, provide them with information on integration with their own tools and finally execute their integrated codes on the NEX supercomputing environment. Because of the diversity of the NEX community, researchers often use different programming languages and development environments. In order to improve users interaction with the system, we propose to develop a set of client libraries that will enable users to access data and execute processing components directly from their environment such as MATLAB or IDL. Finally, in order to enable users to easily share the results of their analysis and research, we will provide a set of migration tools that will publish these results using out web services architecture developed under our previous NASA ACCESS award. Users will be able to select from a number of protocols such as OpenDAP, OGC WMS, OGC WCS and others depending on the needs of their community. The proposed enhancements to NEX will improve access to NASA data, model results and services to our current users as well as the potential NCA community.

Deana Pennington/University of Texas at El Paso
Earth, Life and Semantic Web (ELSeWeb): An Earth observation-Driven, Semantic Web System for Computational Modeling of the Impact of Changing Climate on Ecosystems and Human/Environmental Health Systems

The Earth, Life and Semantic Web (ELSeWeb) project will integrate the NASA-funded Earth Data Analysis Center (EDAC, http://edac.unm.edu/) with an analytical Web Service platform, Lifemapper, (Lifemapper.org) which models potential future species distributions under scenarios of climate change. The integrated system will provide climate change impact scientists streamlined mechanisms for discovering, accessing,

understanding, and using Earth observation data, integrating those data into environmental models, particularly in the context of environmental impacts on human health and disease. It will enable them to conduct more complex, more realistic computational experiments (IF-THEN-ELSE models) that include multiple parameterizations of changing environments in addition to changing climates. The integrated system will take advantage of Web Service and Semantic Web technologies, enabling users to visualize a trace of all of the components and parameters that contributed to the model output, and to share resources and collaborate around these. Both EDAC and Lifemapper will be instrumented to automatically collect semantic provenance data as analyses are conducted. Provenance is a trace of all of the inputs and actions that lead up to a derived model product. Additionally, the Semantic Web approach used to generate provenance will also enable automatic mapping of source data selections to transformation algorithms needed to meet the requirements of Lifemapper. With provenance, users will be able to access and visualize information regarding original Earth data sources; processing and analyses conducted on those data at EDAC; algorithm selection made by the user; model parameters set by the user; and processing conducted within Lifemapper. Relevant provenance information will be automatically captured along with Lifemapper's generation of models, minimizing the need for manually input metadata.

The ELSeWeb platform will enable modeling of climate change impacts on animal and plant species using Earth observation and climate change data, through the integration of these independently successful initiatives. This, in turn, will enable highly innovative new modeling of complex factors associated with biotic change such as health and infectious disease, that depend not only on climate change and species distributions, but also on other human/environmental interactions. Provenance generation capabilities of ELSeWeb services will enable modelers to experiment with many potential future scenarios of human/environmental system, while tracking each computational experiment such that analysts of model output will understand the data, model and parameter choices that were made. These unique capabilities will transform scientists' ability to collaboratively investigate the implications of different climate change scenarios for a wide variety of human/environmental systems, particularly as they relate to the nexus of environmental factors in human health and infectious disease.

Rahul Ramachandran/University of Alabama, Huntsville
Curated Data Albums for Earth Science Case Studies

Case study analysis and climatology studies are common approaches used in Atmospheric Science research. Research based on case studies involves a detailed description of specific weather events using data from different sources, to characterize physical processes in play for a given event. Climatology-based research tends to focus on the representativeness of a given event, by studying the characteristics and distribution of a large number of events. 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. For example, in hurricane science, finding airborne or satellite data

relevant to a given storm requires searching through web pages and data archives. Background information related to damages, deaths, and injuries requires extensive online searches for news reports and official storm summaries.

A team of research scientists and informatics experts proposes to use mature content aggregation technology to create curated "Data Albums" to support case study analysis and climatology studies. We will leverage two established technologies to create these Data Albums: the Noesis search engine for content aggregation with Drupal content management technology providing the underlying framework and album-style user interface. We will use emerging information casting protocols for publishing data and information.

We will use a science-driven approach to enhance the existing Noesis components to create Noesis 2.0, an open source, reusable aggregation tool to compile online science data and information into Data Albums for specific events or topics, to facilitate case study based research. These interactive Data Albums will contain compiled information regarding relevant data for specific events, filtered based on geophysical parameters, geolocation, time and scientific relevance. They will also contain aggregated supplementary information to augment research including news articles, reports, images and videos detailing events and socio-economic impacts, related scholarly articles useful for a background literature survey and other useful information such as weather reports. Data Albums will support curation by the album author, host data center and, after an album is made public, the research community. Such social curation will include annotations and ratings on the contents, with privileged users able to add or remove resources. Data Albums' contents will also be published in XML as information casts to facilitate automated access.

The science drivers for the Noesis 2.0 Data Album concept come from two of NASA's Science Focus Areas: Water and Energy Cycle, which fosters research to improve hurricane prediction, and Weather, seeking improved understanding of the Earth system through better understanding of weather processes. The proposed effort will deploy curated Data Albums for two science research areas. The first instance will be a portal for hurricane case studies at the Global Hydrology and Resource Center (GHRC), a NASA Data Center. This portal will auto-generate Data Albums for specific hurricane events, compiling information from NASA field campaign collections and other sources. The second instance will be customized to support NASA's Short-term Prediction Research and Transition (SPoRT) Center in compiling case studies to evaluate the performance of numerical weather prediction tools for convective weather events.

Installations of Noesis 2.0 Data Albums will remain in operation at GHRC and SPoRT after the end of ACCESS project funding. Noesis 2.0 will be made open source and shared via the Drupal code repository, to encourage reuse of the software, and ensure the continued utility of this tool. The team will also actively promote the tool to colleagues at other NASA Data Centers and within the ESIP Federation.

David Roy/South Dakota State University
Enhanced Web Enabled Landsat Data (WELD) Access

The NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs) funded Web-enabled Landsat Data (WELD) project is systematically generating 30m composited Landsat Enhanced Thematic Mapper Plus (ETM+) mosaics of the conterminous United States and Alaska. The WELD project has demonstrated the potential of large volume Landsat data processing to provide, on a systematic basis, continental scale 30m Landsat data products in support of the Landsat user community. Under current MEaSUREs funding we have developed an intuitive what you see is what you get (WYSIWYG) WELD product Internet distribution interface, developed using open source software, and now operational at the USGS EROS with more than 250 registered users (http://weld.cr.usgs.gov/).

This focused, modestly budgeted, two year ACCESS proposal seeks to undertake research to support commonly articulated outstanding user requirements for (1) provision of the WELD products in different map projections, (2) expansion of the WELD product distribution to global scale, (3) support of Open Geospatial Consortium (OGC) compliant WELD products.

William Teng/Wyle Information Systems. LLC
Bridging the Digital Divide between Discrete and Continuous Space-Time Array Data to Enhance the Accessibility to and Usability of NASA Earth Sciences Data for the Hydrological Community

A longstanding and significant “Digital Divide” in data representation exists between hydrology and climatology and meteorology. Typically, in hydrology, earth surface features are expressed as discrete spatial objects such as watersheds, river reaches, and point observation sites; and time varying data are contained in time series associated with these spatial objects. Long-time histories of data may be associated with a single point or feature in space. In meteorology and climatology, remotely sensed observations and weather and climate model information are expressed as continuous spatial fields, with data sequenced in time from one data file to the next. Hydrology tends to be narrow in space and deep in time, while meteorology and climatology are broad in space and narrow in time.

This Divide has been an obstacle, specifically, between the hydrological community, as represented by the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) and relevant data sets at the Goddard Earth Sciences Data and Information Services Center (GES DISC). CUAHSI has developed the Hydrologic Information System (HIS), which is built on international geospatial standards, with one of its aims to bridge the Divide. GES DISC has a long history of supporting energy and water cycle data sets, represented by two of its discipline-oriented entities, Precipitation DISC (PDISC) and Hydrology DISC (HDISC). The opportunity costs of the Divide are high. It has largely prevented the routine access and use of NASA Earth science data by the hydrological and, more generally, geospatial community.

There have been a number of preliminary efforts, over the past several years, to explore ways to bridge the Divide. One specific example is an HIS Web service providing access to selected MODIS data, via the NASA Giovanni visualization and analysis system. A more recent effort by GES DISC and EPA BASINS (Better Assessment Science Integrating point and Nonpoint Sources) resulted in a customized service enabling the access to 30-year time series of a single parameter of the NLDAS (North American Land Data Assimilation System) data set, based on a reorganization of the archived data and the GrADS Data Server (GDS). An ongoing prototype effort by CUAHSI, GES DISC, and the NASA GSFC Hydrological Sciences Branch is aimed at demonstrating the feasibility of accessing selected NLDAS parameters, via WaterML-compliant Web services and GDS, by HIS users from within a client such as HydroDesktop. All of these prototype efforts have shown the potential for bridging the Divide. Significant impediments remain, however, including the scaling from a few test parameters to a vastly greater number of parameters.

The overall approach of the proposed project comprises the following:

1.Incorporate NASA data into the HIS ontology; integrate NASA data into the HIS OpenGIS Catalogue Service Web (CS-W) specification.

2.Develop WaterML Web services for NASA data; reorganize NASA data for optimal access and use.

3.Enhance the CUAHSI HydroDesktop system to (1) access NASA's continuous array data and summarize them over discrete spatial objects, such as watersheds and (2) enable comparison of time series of point observations from monitoring site locations with those of remotely sensed data measured over the same locations.

4.Develop hydrological use cases to guide the implementation, and serve as the metric for the usefulness, of the project technologies.

This project would represent a significant advance in extending NASA Earth science data to a large hydrological user community that has, heretofore, mostly unable to easily access and use NASA data.

Charlie Zender/University of California at Irvine
Simplifying and Accelerating Model Evaluation by NASA Satellite Data

The fidelity of geoscientific model results are increasingly evaluated by comparison to products derived from NASA satellite measurements. The satellite data are archived in HDF-EOS format, which is now a superset of the netCDF format employed by most geoscientific models. Putting NASA-generated (HDF-EOS) data and model-generated (netCDF) data on a common grid, in the same format, for numerical comparison can be arduous because of data format incompatibilities. Moreover, some analysis tools for netCDF data have no counterparts or equivalents for HDF-EOS data. Many researchers

desire a common toolkit for both HDF-EOS and netCDF data would 1. simplify and accelerate the independent analysis of both data formats (HDF-EOS and netCDF), 2. exploit the strengths of netCDF's underlying HDF data format with easy-to-use netCDF tools, 3. ease evaluations of model predictions (in netCDF format) by NASA-generated data (in HDF-EOS format).

The primary purpose of this project will simplify the workflow involved in intercomparing HDF-EOS format data to model results in netCDF format. It will do so in a user-friendly and transparent way, by improving the netCDF Operators (NCO) which are robust components of the scientific data analysis software stack already employed at most Earth science modeling centers. The key NCO improvement will be to support group hierarchies. Groups are nestable namespaces that allow for hierarchical storage (the ``H'' in HDF). Utilizing groups to store ensembles of observations and predictions would vastly simplify and accelerate the characterization, evaluation, and intercomparison of multiple geophysical observations and simulations. Until now this has been impossible since NCO supports only ``flat'' datasets. The proof of this claim will be demonstrated by applying the improved NCO to a prototypical, NASA-relevant, Earth System Science research problem: to characterize, evaluate, and intercompare Earth System Model-simulated and NASA-retrieved snow cover and albedo trends and variability in the CMIP5 models to be used in the IPCC AR5 climate assessment. NCO is a robust element of the scientific software stack used by the community of Earth Science researchers inside and outside of NASA for over fifteen years. Researchers worldwide employ NCO's user-friendly commands, honed through years of open source, developer-user feedback, to process terascale model datasets (often in preparation for comparison to HDF data). However, there is not yet an NCO-equivalent for processing HDF-EOS data. This is partly because NCO does not yet understand all the powerful HDF capabilities now accessible through netCDF API. The project will remediate much though not all of this deficiency. The primary outcome will be the applicability of NCO to ever-increasing sets of HDF-EOS data, and of netCDF data, that utilize groups to organize and contain data. The proposed work directly responds to the ACCESS call to increase use of EOS data by the climate modeling community analyzing and evaluating the CMIP5 simulations. Moreover, the improved NCO capabilities will apply to all geophysical data archived in HDF-EOS and netCDF formats. The significance of the proposed work is expected to be greatest for applied science researchers wishing to more fully exploit NASA data to evaluate model simulations. The PI is a long-standing climate modeler, software developer, and NASA-funded researcher who understands many of the barriers to model evaluation and who has developed, in the form of NCO, an elegant solution to some of them. The PI participates in the relevant geoscientific communities, including as a reviewer for the ESDS Standards Process Group, for the last two IPCC climate assessments, and in the development of ESG-supported models such as the community Earth System Model