Heliophysics Data Environment Enhancements
Abstracts of selected proposals.
(NNH10ZDA001N-HDEE)

Below are the abstracts of proposals selected for funding for the Heliophysics Data Environment Enhancements program. Principal Investigator (PI) name, institution, and proposal title are also included. 18 proposals were received in response to this opportunity, and 10 were selected for funding.

Richard Bogart/Stanford University
Data Services Continuation of the SOHO/MDI and Stanford Helioseismology Archive

The SOI/MDI data archive is being migrated from its original architecture to the Data Record Management System (DRMS) used by the AIA/HMI joint data archive for the Solar Dynamics Observatory. This will not only provide a stable environment for maintenance of the archive for the duration of the SDO mission, but also provide greatly enhanced capabilities in terms of data search, selection, retrieval, and delivery. In addition to the SOI/MDI archive, other datasets for which the Stanford Solar Observatories have assumed archive responsibility will also be migrated to this system.

The SOI/MDI data are expected to be a vital data resource for the fields of solar and heliospheric physics for many years, as they provide a continuous and consistent baseline of information covering more than a solar cycle before the launch of SDO and much or all of the observing lives of other major solar space missions.

It is probable that the MDI data can and will be improved through through cross-calibration with HMI data. Although activities such as recalibration are beyond the scope of this proposal, if they are undertaken by others, it will be incumbent on the responsible and knowledgeable personnel at the archive to assure the data availability, provide expert assistance to those involved, and incorporate the results in the archive.

A substantial fraction of the existing MDI data archive is for Level 2, post-calibration data resulting from scientific analysis. New results and discoveries as well as new techniques of data analysis resulting from SDO and especially HMI will certainly call for reexamination and new or improved analyses of the long data series from its predecessor. These too will require support and resources if they are to be incorporated in the Resident Archive

We are proposing a basic archive maintenance program for the SOI/MDI data as well as other ground-based data sets of related historical interest that will assure continuing access of the scientific community to the data, along with expertise as needed and requested to help in proper interpretation of the data; support for any changes that may occur in the structure of the archive; periodic validation of the data and its supporting documentation as necessary; support for any recalibration or reprocessing activities that may be deemed useful; and support for the incorporation of newly processed analysis

data sets in the resident archive. Maintenance of the archive in a ready and active state will be assured by its incorporation and use within the SDO DRMS, of which we are the principal developers and maintainers.

Jorg-Micha Jahn/Southwest Research Institute
Value Added Services for VxOs: Deploying an Energetic Neutral Atom Image Inversion Database and Web-Based Inversion Tool

In recent years, space plasma physicists in magnetospheric, planetary, and heliospheric physics have been using a new kind of measurement: Energetic Neutral Atom (ENA) remote sensing of plasma processes, usually performed in the form of 2-D neutral atom imaging. This approach, which relies on the conversion of plasma particles into energetic neutrals via charge exchange, allows us to observe plasmas from a distance, freeing us from the restriction of having to measure plasma in situ and thus locally only. Energetic neutral atom measurements, however, add a new challenge themselves: they need to be mathematically inverted to expose the plasma distributions causing them . Unfortunately, ENA inversions are difficult to create. Even after 10 years of intensive ENA data collection, this remains a very high entry barrier to the meaningful and comprehensive analysis of ENA data.

We propose to overcome this problem at least in Earth's magnetosphere by creating a value-added service to the existing NASA VxO infrastructure. The key objective of this proposal is to provide the magnetospheric research community with easy access to ENA inversions for magnetospheric ENA data. Our main thrusts are: (1) Create easy access to ENA inversions via a user-friendly, web-based service. We will enable a straightforward creation of custom ENA inversions as well as provide a browse interface to a comprehensive database of pre-calculated ENA inversions covering the IMAGE and TWINS missions for ENA energies above 1.0 keV. We integrate this service (and attached database) into the existing Magnetospheric Virtual Observatories (VMO) at GSFC and UCLA. (2) Increase the speed of ENA inversions to create an interactive (i.e., short response time), web-based user experience. We propose to bring easy and fast (!) access to custom ENA inversions to the whole community. (3) Establish and disseminate ENA inversion quality factors. We have developed a set of ENA inversion "quality factors", that, combined with a User's Guide, help uninitiated researchers to quantify and evaluate (!) the quality of an inversion themselves.

The project will provide the following end-products: a database of pre-calculated inversions for IMAGE & TWINS ENA data above 1.0 keV, quality factors for each individual (!) inversion, a computationally fast inversion code, a single and user-friendly web-access tool to allow researchers to interactively create inversions and browse pre-calculated inversion results, as well as a User's Guide to help researchers perform inversion tasks online and to evaluate the results.

Homa Karimabadi/SciberQuest, Inc.
Value Added Services for VxOs: Intelligent Archiving and Automated Physics Mining of Large Data Sets

We propose to develop an intelligent data analysis engine that extends the functionality of Virtual Observatories (VxOs) from data portal to science analysis resource. Heliophysics is a data centric field which relies heavily on the use of spacecraft data for further advances. The prevalent approach to analysis of spacecraft data is based on visual inspection of data. As a result, the vast majority of the collected data from various missions has gone unexplored. While the need for advanced algorithmic approach to data exploration and knowledge discovery is generally recognized by experimentalists, the adoption of such techniques ("data mining") has been slow. This has been partly due to the steep learning curve of some of the techniques and/or the requirement to have a working knowledge of statistics. Another factor is the existence of a plethora of data mining approaches, and it is often a daunting task for a scientist to determine the appropriate technique. Our goal has been to make such tools accessible to non-experts and remove it from gee-whiz domain to a practical tool that will become part of the standard arsenal of data analysis. To this end, we have developed an automated data mining technique called MineTool. In a short time, MineTool has gained significant traction among experimentalists. There are currently several ongoing data analysis projects involving MineTool, including three recently funded NASA grants, in magnetospheric and solar physics. Event detection is perhaps the most familiar example of data mining application to a non-specialist. However, data mining offers much more versatility and can be used for (i) classification, (ii) dependency analysis, (iii) modeling, (iv) statistical analysis, and (v) anomaly detection.

The goal of this proposal is to leverage our recent success and offer a complete solution for data analysis as value-added service to VxOs. The proposed tasks are: i) Development of automated data preparation algorithms, ii) Integration of data preparation and data mining algorithms with VMO/VHO, iii) Use of our group of beta testers, experimental space physicists, to obtain feedback on the ease of use and functionality of the service and incorporation of the feedback into the design process, and iv) A real life deployment of the software for creation of event lists by our collaborators. These innovations will significantly enhance the science return from NASA missions by providing data centers and individual researchers alike an unprecedented capability to mine vast quantities of data.

KD Leka/NorthWest Research Associates, Inc.
Data Services Continuation: The Imaging Vector Magnetograph Resident Archive.

The goal of this "Data Services Continuation" proposal is to establish the Resident Archive for data from the Imaging Vector Magnetograph ("IVM") at the U. Hawai`i/Mees Solar Observatory. In synoptic operation 1991--2006 with standard observing programs that included ``movie-mode'' magnetogram time series, data from the IVM have proved extremely scientifically valuable, yet their availability to the broad community has been very limited.

The proposed effort is a follow-on to our present Data Services Upgrade effort (NASA/VxO contract \#NNH09CF83C to NWRA). In the single year of funding, we have upgraded and streamlined the IVM data reduction pipeline, tested and chosen code to perform the inversions and ambiguity-resolution, acquired the entire collection of raw tapes/DVDs from U. Hawai`i, recovered three years' worth of Level-0 data from aging tapes, initiated an off-size Level-0 data repository, and begun building a searchable data-hosting website at NWRA.

As a Resident Archive, almost a solar-cycle's worth of active region vector magnetograms will be available on-line, searchable and retrievable through the Virtual Solar Observatory. Scientifically quantitative measurements will be delivered in physically meaningful units, including estimates of measurement uncertainties, using very generic FITS files. Metadata which are searchable by the VSO will be automatically produced. The IVM Resident Archive will be supported by an expert on these data and their interpretation, and accompanied by a suite of analysis software and full documentation. The availability of this database will allow the Solar Physics community to perform research directly related to solar-cycle active region characterization, energetic event initiation, and other topics in support of NASA's Science Mission Directorate.

Daniel Morrison/JHU/APL
Multi-satellite Instrument Data Set Conjunction Locator Service: Value Added Services for VxOs

We propose to develop a series of light weight web services for use by the Virtual Ionosphere-Mesosphere-Thermosphere Observatory (VITMO http://vitmo.jhuapl.edu) and other VxOs that allows the overlap between multiple satellite data sets to be determined, allowing the VxO to supply both sets of overlapping data products. These web services will also allow "near misses", where products are only close in time and/or geographical overlap to be optionally selected by the user based on criteria that the user provides. This proposal is focused on providing a generalized set of services that will initially support coincidences between the SABER, TIDI, and GUVI instruments on the TIMED satellite, the SOFIE and CIPS instruments on the AIM satellite, the SUSSI instruments on DMSP F16, F17, and F18, as well as C/NOFS and COSMIC satellites. This proposal will also develop a sufficiently general system that can be easily extended to support additional satellites and instruments. This proposal will add a "Value-added Services that incorporate significant enhancements to the services provided by one or more of the VOs". These services are based on the existing coincidence service in VITMO that calculates satellite data set and ground site coincidences for various remote-sensing instruments. The service will be exposed to other VxOs as a web service (with WSDL and REST interfaces) with the outputs as an event list compatible with the SPASE HELM program. The proposed services will allow the non-specialist user to select data that they were previously unable to locate, opening up analysis opportunities beyond the instrument teams and making it much easier for future students who come into the field.

This value added search service is an excellent follow-on to the existing search strategies in the Virtual ITM Observatory.

Stuart Nylund/JHU/APL
Data Services Continuation: Resident Archive for Energetic Particles and Ion Composition (EPIC) Instrument on Geotail

This proposal seeks funding to continue the data services of the Geotail/Energetic Particles and Ion Composition (EPIC) instrument as a Resident Archive. In anticipation of NASA's decision to discontinue funding for both Geotail flight operations and the EPIC Science Team, we are proposing 1) to continue serving the valuable EPIC data set to the community and providing our expertise in the interpretation of EPIC data, 2) to preserve the integrity of the EPIC data set and maintain the data services by which these data are accessed, and 3) to prepare the EPIC data set for a transition to a Final Archive.

James Ryan/University of New Hampshire
SMM/GRS Data Preservation and Publication

Level 1 data from the Gamma Ray Spectrometer Experiment on the Solar Maximum Mission now reside on digital video tapes at three institutions accessible only on outdated VMS machines. These data represent the only solar γ-ray observations that completely cover the period of 1980 to 1989. The measurements they represent revolutionized our concepts of high-energy solar physics. They spawned other high-energy solar physics missions, both domestically and internationally. However, they were not widely disseminated in their raw form, primarily because of the limited technology at the time, subtleties of the evolving instrument response and the subtleties of background subtraction techniques. However, since then, through much work and understanding, the instrument response has reached a stable state as have the background subtraction techniques that have been used elsewhere with success. These facts plus the volatile nature of the data media present an opportunity to the solar physics community, that being that we are in a position to (1) preserve the data in a standard format on long-lived media, (2) post the data on a public forum, such as the Solar Data Analysis Center (SDAC) at Goddard Space Flight Center and (3) use facilities such as the Virtual Solar Observatory (VSO) to serve the data to the solar physics community so that they can access, interpret and publish these data.

To that end, we propose to port the data from the volatile tape media to longer-lived media at the SDAC (http://umbra.nascom.nasa.gov/) in a form that lends itself to access and manipulation by the scientific community, with file reading tools that can be integrated into the SolarSoft system.

Robert Stockwell/NorthWest Research Associates
Data Services Upgrade: Enhancement of WINDII Measurements for the Virtual Observatory

The Wind Imaging Interferometer (WINDII) aboard the Upper Atmosphere Research Satellite (UARS) measured winds and temperatures in the 90-270 km range between 1991 and 1997. WINDII remains the only source of the global wind and temperature fields between 120-270 km. These data have remained relatively unexploited, yet they have very high value for studies of thermosphere-ionosphere coupling, and solar effects on the dynamics of the upper atmosphere. We propose to incorporate the entire WINDII measurement dataset (both Level 3 data and Level 2 data) into the NASA's Space Physics Data Facility (SPDF) and the Virtual ITM (Ionosphere Thermosphere Mesosphere) Observatory (VITMO) at JHU/APL. We will employ the CDAWeb (Coordinated Data Analysis Web) system, which has an extensive set of tools and a support desk for users.

At the present time, only Level 3 wind and temperature data are available online, stored in nonstandard Unix binary files. We will develop the software to process this dataset along with Level 2 data (including "line of sight" data and volume emission rate profiles), and enhance the metadata by adding orbital information and data quality measures. The datasets will be processed into the formats required by the SPDF and VITMO. These systems will automatically provide enhanced data search services, a single data delivery mechanism, support for data imagery and "Search for Data by Browsing Images" (SDBI), and ensure the data are findable and appropriately usable by the community through the Heliophysics virtual observatories. The data will be accessible in a variety of platform independent formats and provided alongside other similar measurements, thus providing simple access to researchers using TIDI, HRDI and other instruments. The result of this project will be to improve the quality, utility, and accessibility of datasets relevant to Heliophysics research.

Our efforts will greatly enhance the use and utility of the WINDII measurements. WINDII has helped and continues to help advance the science of Solar Terrestrial Relations, and therefore our understanding of the coupling between the lower and upper atmosphere. This project addresses Goals 3A and 3B articulated in the NASA strategic plan [from NASA's Fiscal Year 2009 Performance and Accountability Report (PAR)]: Study Earth from space to advance scientific understanding and meet societal needs (3A); and Understand the Sun and its effects on Earth and the solar system (3B). The proposed effort also directly addresses the HP Data Policy [NASA Heliophysics Science Data Management Policy Version 1.1, 12 April 2009]: "Spaceflight projects ("missions") are the core of the Heliophysics Data Environment (HPDE). The data from missions is to be made available by them both directly and via Virtual Observatories (VOs) that will provide one-stop access to data from many missions along with tools for cross-mission analysis and visualization."

Robert Weigel/George Mason University
Value Added Services for VxOs: An API and Server Software for Merge, Subset, and Filtering of Time Series-Like Data

We propose to (1) develop a standard API for time series-like data, (2) develop a software package, TSDS (Time Series Data Server), that implements this API that provides server-side super-setting, sub-setting, filtering, and uniform gridding of time series-like data, (3) make data available through TSDS on several key servers in the heliophysics environment, and (4) develop client-side software for standard data analysis packages (IDL, MATLAB, Java, Python, and Excel) that will allow access to a TSDS-enabled server.

Yongliang Zhang/Johns Hopkins University / APL
Data Service upgrade: Storm-time global thermospheric O/N_2 ratio from IMAGE FUV data

We propose to derive unique and high cadence global thermospheric O/N_2 ratio using NASA IMAGE FUV (SI-13) data during geomagnetic storms between 2000 and 2005, saved the O/N_2 data, their associated UT, date, latitude and longitude in NetCDF files, post the NetCDF files at a permanent TIMED/GUVI website (http://guvi.jhuapl.edu) for public to access. We also provide all metadata information for VITMO (Virtual Observatory of Ionosphere, thermosphere and Mesosphere) to access to the dataset. Storm-time O/N_2 change (depletion and enhancement) is one of the key space weather metrics. It directly impacts the ionosphere condition, magnetosphere-ionosphere coupling and satellite drag. The O/N_2 from TIMED/GUVI has been widely used in the thermosphere and ionosphere research (storm time, season effect) and served as validation source for the global simulation (e.g TIMEGCM). However, due to long re-visit time (97 min), GUVI O/N_2 is not suitable to trace the fast thermosphere variation during storm times. The IMAGE FUV data has 2 min time resolution and covers half of the earth most time, the IMAGE based O/N_2 data will help to provide new insight to the fast and global thermosphere responses during storm time. Furthermore, the proposed method has been proven to be robust and it can produce reliable IMAGE O/N_2 without the requirement of accurate calibration of IMAGE FUV data. It just requires that the IMAGE FUV sensitivity is stable over a period of a few days (typical lifetime of magnetic storms). All IMAGE FUV data (2000-2005) are stored locally at Johns Hopkins University Applied Physics Laboratory (JHU/APL). They can also be retrieved through the NASA database at http://cdaweb.gsfc.nasa.gov/. The unique O/N_2 data will allow the community to study the fast response of the thermosphere to the storms thus improve space weather (ionosphere) forecast, and serve as validation sources for global ionosphere and thermosphere simulations.
This proposal directly addresses "Data Service Upgrade", one of the goals of NASA's Heliophysics Data Environment Enhancements program. The resulting data set also significantly benefits the researches that address NASA's objective (3B): "Understand the Sun and its effect on Earth and solar system". This proposal fits the key requirement and the focus areas of the Geospace Science Program: "To understand the space that surrounds and is influenced by the various solar system bodies, including the ionosphere,

the magnetosphere and beyond" and "Open the Frontier to Space Environment Prediction".