

Update on HDMC activities, 3 May 2010
Aaron Roberts, HDMC Project Scientist

While we have not had any group meetings for some time, many things have been going on to advance the data environment in various ways. Note that the next general HDMC meeting will be 5-6 August 2010 at GMU in conjunction with the next Informatics workshop that Tom Narock and Bob Weigel will be hosting 2-4 August.

Central to our efforts is the SPASE Data Model progress and stability. It was satisfying to be able to assert to the HP extended missions Senior Review (which took place mid-April) that the Data Model is, in fact, stable, and this certainly makes life easier. Continuing tweaks are occurring, but these will not cause problems for existing resource descriptions. Many such descriptions are being created, and we're making good progress toward completeness. This is tedious work, and it will take perseverance. Collaboration with HELIO in the Data Model and other areas has begun, and needs to be improved.

In the mission Senior Review context, the projects had to submit Mission Archive Plans, which were reviewed separately for input to the Panel report. The general good news that can be reported from this is that all the missions are now doing a good to excellent job of making data available in useable forms with informative documentation. The open data policy is working. The remaining major problems, for some missions, have to do with working out final archiving arrangements and producing SPASE descriptions to integrate the datasets into the VO and/or cross-VO framework. Plans are moving forward on these problems (see the next item, in part).

The HP Data Center Senior Review process is now complete (see: <http://science.nasa.gov/heliophysics/heliophysics-data-centers/>), and HDMC is officially moving forward. There were no specific recommendations for changing the HDMC plan, so Jeff Hayes has said we should execute what we proposed. The funding will not go up, but so far we have been assured a level budget, which should be enough to do good things. A significant concern is that the report notes that there does not seem to be a groundswell of support for integrated access to resources. We have our work cut out to demonstrate relevance. Successful use cases and at least a few ardent supporters who see our tools as essential are important for continued progress. Please tell me about any such uses or people.

The main logistic HDMC problem is getting the VxOs and related activities funded on a new basis as the NRA funding ends. I know the relevant VxO groups are working on proposals consistent with what we said we'd do, as I requested a while back. Hopefully these will go smoothly through the system; the route seems at least fairly clear. It somewhat surprises me that most groups have not responded yet to my request for short proposals. Let me know if there are issues that need to be resolved or any other questions.

The Heliophysics Data Environment Enhancements (HDEE, the successor to the "VxO" call) selections were made a few months back, and the abstracts are now available. See:

<http://nspires.nasaprs.com/external/viewrepositorydocument/cmdocumentid=227199/HDEE09%20Selections.pdf>

or go to the nspires.nasaprs.com website and navigate from the "Selected Proposals" link in the upper left corner to the HDEE section. Nearly all of the grants and contracts are now in place. There were no new VxOs called for, and therefore none selected, so the bulk of the effort will go into Data Upgrades and Resident Archives, with a couple of significant Value Added projects.

On more specific items, in particular the working groups we had identified, registries are essential to the HPDE, and we now have "git" repositories to keep track of resources and their provenance. In addition, Todd and others have developed tools to use these registries, and there are demos of the capabilities. Todd recently sent out a note about the availability of open source projects in this area. VSPO is continuing its work on completing a "header level" inventory of HP resources, and to swap in VxO provided more extensive descriptions as they are available. It will be important to come to a clear resolution on the architecture of the registries in the near future so that others will more easily be able to access the resources the VxOs and others are making available. Related to this, efforts to make SPASE-QL a cross-VxO reality are continuing. The current effort is devoted to providing uniform basic access to metadata and data. The Data Center Senior Review was strongly supportive of this effort.

AutoPlot continues to be developed, with the improvement of the usability and robustness being high priorities. This effort will continue to be funded as an identified project. This will be our major "cross disciplinary browser," at least for time series data. It's ability to accept and output many formats will be useful in itself. The support of some RBSP groups for AutoPlot development should lead to both better capabilities and wider adoption. ViSBARD is receiving a trickle from HDMC and some other support to keep the 3-D data visualization efforts alive, and it has recently added the ability to include views of simulations along with the data. Tools such as these need to become really easy to use and more widely known if they are to ever have much impact.

The ability of AutoPlot to read and write many formats is related to the "DataShop" efforts to solve the formats problem. The latter work will now be funded under the HDEE NRA, but to be successful, all the VxOs should be included at some level in the effort. Now that we are getting the basic metadata and data access problem pretty well in hand, the next problem is to provide users with data they can use without wasted effort, and this project should make this possible.

I'd like to reinstitute telecons on an approximately biweekly basis, probably starting May 12, and to get the working groups on the above topics (Registries, Interoperability, Visualization, and Data Processing Services) moving more systematically. Efforts in all these areas are proceeding, but need to be coordinated better across the various groups.