

IN23B-1358



VOs and Heliophysics: Would anyone like some CASSIS?

Bentley, R D (Dept. Space and Climate Physics, University College London, Dorking, United Kingdom; rdb@mssl.ucl.ac.uk), **Lapenta, G** (Plasma-astrophysics Section, Katholieke Universiteit Leuven, Leuven, Belgium), **Blanc, M** (Centre for the Study of Radiation in Space, National Centre for Scientific Research, Paris, France), **Fox, P** (Tetherless World, Rensselaer Polytechnic Institute, Troy, NY, USA), **Walker, R J** (Institute of Geophysics and Planetary Physics, UCLA, Los Angeles, CA, USA) and The CASSIS Team

Virtual Observatories related to heliophysics have emerged in various regions of the World. Partly because of their sources of funding, the emphasis they place on aspects of the problem differ and how well they can be integrated with other capabilities varies to a greater or lesser degree. If science within the Solar System is to address more diverse Use Cases we need to be more proactive in our efforts to achieve interoperability – **CASSIS** is designed to do this.

A large number of virtual observatories have been developed within NASA's *Heliophysics Science Division* – they include: the Virtual Heliospheric Observatory (VHO), Virtual Space Physics Observatory (VSPO), Virtual Magnetospheric Observatory (VMO), etc. In addition, the Virtual Solar Terrestrial Observatory (VSTO) is funded by NSF. These VxOs deal mainly with the Sun-Earth Connection – currently the only virtual observatory related to planetary data is Europlanet RI, funded by the EC under FP7.

The general compliance of the NASA VxOS with the SPASE data model means that they are interoperable to an extent, but it is not uniformly applied across all domains. Planetary data supplied by NASA's Planetary Data System is changing to handle PDAP but this will take time. The SPASE data model needs to be part of a bigger picture to ensure that Solar System science continues to develop.

CASSIS, the *Coordination Action for the Integration of Solar System Infrastructure and Science*, is intended to facilitate the discussions that will lead to the level of interoperability that is required to support investigations across all parts of the Solar System science. **CASSIS** is looking at how we can make data more accessible and interoperable, how metadata can be made more universal in its usefulness, and how services can be used more widely.

We want to do this in a way that is attractive to the users and is not onerous.

Making data more accessible and interoperable

Small changes to the way that data are stored can make them more accessible and improve interoperability:

File Names – There are no hard and fast rules, but the name needs to be sufficiently unique that:

- The type and origin of the file can easily be identified (time of observation?)
- The file can exist without causing confusion when removed from the context of where it is normally stored

Directory Structure – A hierarchical directory structure makes it easier to find files (and is strongly preferred)

- Ideally the structure should be a tree based on dates
- Essential for resource-poor providers; beneficial for a data centre

Summary of Observations – It simplifies access if the archive maintains a summary of the observations that have been made

- Particularly useful if all the observations are not available on-line

Standards and Interoperability

Drive for interoperability comes from increasing desire to undertake cross-disciplinary studies. **HELIO**, for example, is a Virtual Observatory that is trying to facilitate this – domains involved include solar, heliospheric, planetary, geophysics...

To achieve its objectives, **HELIO** must address many issues – these are the things that make everything possible...

Must support a search across several domains

- Search based on metadata and derived products
- Search related to phenomena evolving in 4-Dimensions
- Need to condition the metadata to ensure that temporal and spatial coordinates are homogeneous and interoperable

Must provide integrated access to data from many domains

- Different file formats, ways of storing, handling and using the data
- Variety of access techniques & protocols (http, ftp, Web-services...)
- Accessibility depends on how providers have organized things

HELIO is developing data models to describe the heterogeneous systems:

- Single data model is not realistic option – trying to create a model that over-spans the others based on models from EGSO, SPASE, IVOA, etc.
- Data model is important since we **need to annotate the metadata** to properly describe quantities
- Unambiguously define the parameters involved and puts them in context of other information that are used

Developing recommendation for providers about how to name and storage strategies that would make their data archives more interoperable

Developing standards for files that **HELIO** is producing

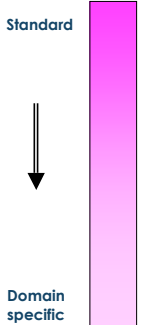
- Include suggested sets of quantities, etc. that should be used
- Annotation also an important part of this activity
- If adopted elsewhere, ideas engender interoperability

File Metadata

It is essential that all files contain good metadata describing the characteristics of the observations

- If metadata are not properly formed, it may be impossible to use the data in some circumstance
- Files should be in some standard format
 - FITS, CDF or similar
- Need to try to push down the boundary between the standard and domain specific areas of the metadata
- Some form of annotation would be beneficial

The key is increasing the part of the metadata that is standard



Want more Information?

Visit the HELIO Web site at www.cassis-vo.eu

