Enabling Interoperability in Heliophysical Domains

Bentley, R D (Dept. Space and Climate Physics, University College London, Dorking, United Kingdom; rdb@mssl.ucl.ac.uk)) 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 examining **HELIO** to find ways in which interoperability could be improved between the services of **HELIO** and in general. **CASSIS** 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.

HELIO has been implemented according to a ***Service-Oriented Architecture***; in this the required capabilities are established as stand-alone services that can be used individually or combined in a workflow. The services provide the ability to identify interesting events or phenomena, determine which instruments could have made relevant observations and then locate and retrieve the required data.

By using a **service-oriented architecture** and adopting **standards developed by international bodies**, **HELIO** is trying to **ensure that its services are interoperable** with those developed by other projects and organizations.

The **HELIO** services have Web Service interfaces based on standards developed by the ***International Virtual Observatory Alliance*** (**IVOA)**. Most services use the **HELIO Query Interface** (HQI) that provides both SOAP and REST like interfaces for normal and long running queries; a few that involved on-the-fly processing follow the asynchronous Universal Worker Service (**UWS**) pattern. Output are generally in **VOTable** format, a standards developed by the **IVOA**; the advantage of **VOTable** is that the parameter can be fully annotated with ***UCDs*** and ***utypes***.

Making data more accessible and interoperable

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

File Types – No restriction except that files should be use a standard format (FITS, CDF, etc.) rather than a proprietary format

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 – the domains involve include solar, heliospheric, planetary, geophysics…

To achieve its objectives, **HELIO** must address many issues – these are the things 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

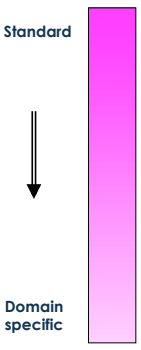
Need **Data Models to describe the heterogeneous syst**ems:
- Single data model is not realistic option – HELIO is creating a model that over-spans the others based on models from EGSO, SPASE, IVOA, etc.
- Data model is important since we ***need to annotate the metadata*** to properly describe quantities

These are generic issues that apply to all Virtual Observatories and agreeing on and adopting a set of standards would make the services of the VOs more interoperable.
Everyone would gain! We would have a growing set of capabilities that could be used a toolkit. The services mixed and matched as needed and we could move forward towards a greater whole.

File Metadata

It is **essential that all files contain good metadata** describing the characteristics of the observations:
- If metadata are not **properly formed**, in some circumstance it may be impossible to use the data
- Some form of **annotation** would be beneficial – this should unambiguously defines the parameters involved and puts them in context of other information that are used

The metadata in files from **different domains share need for certain types of information** – date and time of observation, location of observatory, organization, etc. The **key** is ***increasing the part of the metadata that is standard***.
- Need to try to push down the boundary between the standard and domain specific areas of the metadata



Want more Information?
Visit the CASSIS Web site at ***www.cassis-vo.eu***