CASSIS – Considerations for Collaborative Environments

R.D. Bentley (University College London, UK) and the CASSIS Team

Many aspects of Solar System science are related and there is an increasing desire to address science problems that span disciplinary boundaries. The advent of the Internet and other advances in technology mean that providing access to data is no longer a major issue, however, combining and comparing observations from the different domains is more difficult that it should be, partly because the communities have evolved independently without any consideration of the need for interoperability.

There are several mission and observatories that are under construction – for example, ESA's *Solar Orbiter* mission and the US *Advanced Technology Solar Telescope* (ATST). The projects are starting to think about how the data that they will produce will be managed, how they will be handled and used both internally and in a wider research environment. One way of doing this is by creating a set of services that handle parts of the task and then using these together in a **Collaborative Environment**.

By adopting best practices and following some simple rules, it is possible to improve the quality of the data and make them easier to use by groups not associated with the project. Extending these principles to capabilities that need to be developed to facilitate the project means that they will become interoperable with existing capabilities and enrich the environment available to the use community.

CASSIS has been examining ways in which interoperability could be improved. The project is facilitating the discussions that will lead to the level of interoperability that is required to support investigations across all parts of the Solar System science.

Standards and Interoperability

Solar Orbiter provides an example of the increasing desire to undertake cross-disciplinary studies. To achieve this in an optimal manner we must address many issues…

Must be able to 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 systems:
- Single data model is not realistic option – for example, HELIO using a model that 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 data systems and Virtual Observatories and agreeing on and adopting a set of standards that are also compliant with **IVOA** recommendations would make the services more interoperable and generally useful.

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.

The changes to existing services need not be intrusive – it is often possible to have a new compliant interface co-existing side by side with the current interface.

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

Standard

Domain specific
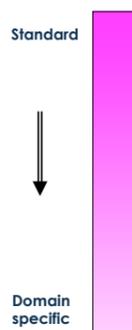
Making data more accessible and interoperable

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

What is CASSIS

- **CASSIS** is a *Coordination Action* funded under Research Infrastructures within the Capacities programme of EC's Seventh Framework Programme (FP7). It started 1 June 2010 with a duration 36 months.
- **CASSIS** intended to facilitate science within the Solar System by improving the interoperability between data and services in all domains
- **CASSIS** brings together three FP7 projects that are directly relevant to this issue – **HELIO**, **EuroPlanet RI** and **SOTERIA**.
Our desire is to engage as many other groups as possible in the discussions, from Europe and the rest of the world

CASSIS is all about interoperability!

Solar System Science has traditionally been undertaken within a number of separate disciplines. Even though many aspects of the system are inter-related it is difficult to address them because of the lack of the integrating tools and techniques. Advances in technology means that this area should not cause a problem

The three projects involved in **CASSIS** have already made progress
- Each is making significant improvements to the infrastructure that supports their communities, increasing ability to do science
- Necessary to coordinate the efforts of these and other projects in order to help break down the inter-disciplinary boundaries barriers

CASSIS trying to go to the next level by cooperating in a number of areas, enabling new combinations of interdisciplinary studies

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