Using HELIO to address multi-spacecraft science: Use Cases and the importance of CASSIS

R. D. Bentley
and the HELIO Team

(RAS-NAM, Llandudno, 20 April 2011)

HELIO, the Heliophysics Integrated Observatory, is a Research Infrastructure funded under EC’s FP7 Capacities Specific Programme

- Started 1 June 2009, duration 36 months
- Consortium includes 13 partners from 7 countries

HELIO will provide the heliophysics research community with an integrated e-Infrastructure that has no equivalent anywhere else

- HELIO will provide the ability to identify interesting phenomena and access relevant solar and heliospheric data together with related magnetospheric and ionospheric data (for planets with magnetic fields and/or atmospheres)
- Need for capability driven by desire to study problems that span disciplines
- Search base on metadata increasingly important as data volumes increase

HELIO will address its challenges following the Integrated Infrastructure Initiative (I3) activities model of the EC’s Framework Programme:

- Networking Activities used to involve the community
- Service Activities used to implement structure of the virtual observatory
- Research Activities used to investigate/develop required capabilities
Science in Context

- Heliophysics, an event-driven science
  - Something is observed and desire is to trace origins or subsequent effects

- Nature of effect depends on causal phenomenon, type of emission, and the location of the observer
  - Most effects have origins in emissions from solar activity; some related to propagating phenomena
  - Location of observer in relation to the source and with respect to a planet determines what is observed
  - Immediate and delayed effects result from the different types of emission

- Presence of magnetic field and/or atmosphere influences effect on planetary environment

- Study requires an understanding of how phenomena evolve in space and time – how they propagate, interact...

---

HELIO addressing a Generic Problem

- Identify interesting things to study
  - Search undertaken in 4-Dimensions across several domains
    - Effects occur as phenomena propagate – whether, where and when to look
    - Follow phenomena through coordinate systems as they evolve
  - Search based solely on metadata and derived products
    - Event lists and feature lists from many domains used as a primary selection criteria

- Review availability of suitable observations
  - Determine whether suitable instruments are at the relevant locations
    - Science objectives dictate types of observations required
  - Determine whether the instruments were making observations
    - Coverage and quality of observations are addition selection criteria

- Locate, select and retrieve the required observations
  - For all domains, system knows which types of data are held where and handles access no matter how data are stored (access protocols & formats)

- Analysis done with users own software tools (e.g. IDL)
Service Oriented Architecture

Design of HELIO splits the tasks into a set of components or services

- Services to aid search process
  - Selection refined based on metadata relevant to the science use case
  - Turn science objectives into required instruments

- Services to locate and retrieve data when selection process complete

- The services can be used independently or as part of a workflow

---

Defined Services

HELIO’s capabilities are defined as a set of principal services that address these tasks

There are also a number of Enabling Services that provide capabilities such as processing, storage, security, etc.

The HELIO Web pages will provide more information as services are developed: http://www.helio-vo.eu/

---

<table>
<thead>
<tr>
<th>Service Name</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search Metadata</td>
<td>Maintain and provide access to existing event data from all domains</td>
</tr>
<tr>
<td>Heliophysics Event Catalogue (HEC)</td>
<td>Maintain and provide access to existing event data from all domains</td>
</tr>
<tr>
<td>Heliophysics Feature Catalogue (HFC)</td>
<td>Maintain and provide access to existing feature data from all domains</td>
</tr>
<tr>
<td>MetaData Evaluation Service (MDES)</td>
<td>Allows the user to create an auxiliary event list based on a newly-derived parameter, etc.</td>
</tr>
<tr>
<td>Context Service (CXS)</td>
<td>Provide context information to help the user make a selection</td>
</tr>
<tr>
<td>Instrument Capabilities Service (ICS)</td>
<td>Match required observation type to one or more instruments (each part of an observatory)</td>
</tr>
<tr>
<td>Instrument Location Service (ILS)</td>
<td>Determine the location of an instrument (part of an observatory) at a specified time</td>
</tr>
<tr>
<td>Observation Coverage Service (OCS)</td>
<td>Provide information on whether an instrument was making suitable observations at a specified time</td>
</tr>
<tr>
<td>Locate and Retrieve the Data (LRAS)</td>
<td>Provide integrated access to data archives in all domains no matter how the data are stored or accessed</td>
</tr>
</tbody>
</table>

Enabling Services

- HELIO Registry Service (HRS)
- Community Interaction Service (CIS)
- Processing Service
- Storage Service
- Auxiliary Information Service (AIS)
- Coordinate Transformation Service (CTS)
- Semantic Mapping Service (SMS)
- HELIO Monitoring Service (HMS)
- Resource Usage Service (RUS)

The HELIO Web pages will provide more information as services are developed: http://www.helio-vo.eu/
Search Process

- **Objective is to turn a science use case that is looking at aspects of a phenomenon into sets of observations of the required type at specific locations and times**

- **HELIO tries to provide as diverse a set of tools as possible, based on metadata from all relevant domains**
  - Heliophysics Event Catalogue (HEC)
    - Catalogue of existing event data
  - Heliophysics Feature Catalogue (HFC)
  - Metadata Evaluation Service (MDES)
    - Allow user to evaluate time series data and generate own event list
  - Propagation, etc. models
    - Helps determine whether, where and when to look
  - Context Service (CXS)

- **How the tools will be use – which and in what sequence – depends on the science problem and the scientist**

Data selection and retrieval

- **Objective is to turn a set of required types of observations at specific locations and times into data from specific instruments at specific times**

- **Some services match type of observations to instruments**
  - Instrument Capabilities Service (ICS)
    - Type of observation each instrument is capable of
  - Instrument Location Service (ILS)
    - Locations of the observatories
  - Observation Coverage Service (OCS)
    - Observation Coverage Table determines if observatory active
    - Unified Observing Catalogue used to handle special cases

- **Once this is done the user should then be able to retrieve the observations they wish to use**
  - Data Provider Access Service (DPAS)
    - Location and method of retrieval described in Provider Access Table
Use Cases and Capabilities

- Let’s examine various use cases and discuss how the search capabilities of HELIO can be used to address them
  - Basic Use Cases
  - Use Cases – Based on Events
  - Use Cases – Propagating Phenomena
  - Instrument Capability and Location Services (ICS & ILS)
  - Heliophysics Event Catalogue (HEC)
  - Metadata Evaluation Service (MDES)
  - Heliophysics Feature Catalogue (HFC)
  - Ability to run and incorporate propagation models

Basic Use Cases

- User knows the time intervals and instruments they are interested in and just wants the data
  - Inputs information through user interface or provides as table
    - VOTable normally used, other formats may be possible
  - Data Provider Access Service (DPAS) finds the data
  - User can selectively retrieve observations based on results from DPAS [and optionally process them]

- User knows the time intervals and locations, and required type of observations, but not the instruments
  - Capabilities Service (ICS) used to identify suitable instruments
    - Also approximate location if they are on/near a planet
  - Location Service (ILS) is required to give more precise location
    - Location of in-situ monitors within planetary environment
    - Observatory moving through interplanetary space
  - Back to basic use case...
Use Cases – Event-driven Queries

- Event-drive queries based on many criteria are possible:
  - Occurrence of a solar flare – timing could be energy dependant
  - Onset of a Coronal Mass Ejection (CME)
  - Energetic particle event
    o Seen at earth; seen elsewhere
  - Disturbance in the solar wind
    o Seen at Earth – (geo) magnetic storm
    o Seen in another planetary environment

- User employs components of HELIO’s search capability to identify time intervals, etc. that they wish to study
  - Uses metadata and associated tools to the identify interval
    o Event and Feature Catalogues (HEC and HFC), Metadata Evaluation Service (MDES) [and Context Service (CXS)]
    o Search may not follow a pre-defined workflow
  - Results can be used as input in search for useful observations
    o Information can be stored on the system if user has authenticated

Identifying suitable instruments

- The instrument capabilities and their observing platforms are described in the Instrument Capabilities Service (ICS)
- The Instrument Location Service (ILS) knows where the observatories are located
- Modelling, by the user or HELIO, identifies whether an observatory lies where it will be affected by propagating phenomena
Use Cases – Propagating Phenomena

- User has identified an event in one part of the solar system and needs related observations.
- Simple Parker Spiral model indicates which observatories are on same or close field lines.
  - User employs model to determine delays related to effects.
- ICS/ILS and OCS determine which instruments probably made observations.
- DPAS used to locate and retrieve the required observations.

Case: Messenger fly-by of Mercury on 29 September 2009

Use Cases – Propagating Phenomena

- User wishes to examine effects of a CME as it propagates through the Solar System.
- HEC used to identify time and location of events on the Sun.
- Propagation model used to determine times & locations in other parts of the Solar System.
- ICS/ILS identify instruments that were suitably located to make the required observations.
- Search tools used to identify other effects elsewhere.
  - HEC identifies major effects in planetary environments.
  - MDES used to look for more subtle effects.
- DPAS locates and retrieves data.
  - OCT used as additional filter.
Use Cases – Reverse Propagation

- User knows of a space weather effect at the Earth and wants to examine the possible
- Propagation model used to determine times & locations back at the Sun
- Search tools aid search
  - HEC used to identify events
  - HFC used to identify relevant features – filaments, CHs
  - MDES used to look for signatures related to shock fronts, etc.
- ICS/ILS identify instruments that were suitably located to make the required observations
- DPAS locates and retrieves data
  - OCT used as additional filter

Services – Status

- Working versions of most services now available
  - Heliophysics Event Catalogue has ~40 lists and growing
  - Heliophysics Feature Catalogue contains details of half a dozen solar and heliospheric features
  - Data Provider Access Service can provide access to over 160 instruments from more than 45 observatories
    - Solar, heliospheric, planetary, geophysics...
  - ICS and ILS working well but evolving
  - MDES in process of being integrated
- Some evolution of services as we try to use them
- Interested in user feedback to help us refine the services
Three ways to access HELIO

- Using the services individually
  - Each service has a GUI; each can also be called through Web Service and similar interfaces from external utilities

- Through the HELIO Front-End GUI
  - The user interface allows the user to combine the services to address a science use case

- From SolarSoft (IDL)
  - Growing number of routines available in a HELIO branch of the SSW tree that integrate the Services into capabilities
  - Currently install as an upgrade (also need EGSO branch)

- Find out more in the Services section of the HELIO Web pages
  - http://www.helio-vo.eu/services
The HEC now contains 40 event lists from various sources. We are experimenting with ways of helping the user identify exactly which list they need for their science objectives.
HELIOS can also be accessed from Solarsoft

HELIOS SolarSoft Crib Sheet

Below are some notes on how to use the IDL code developed for HELIO. The code is based on SolarSoft and at least the HELIO and EGSO branches of the SSW tree must be installed in order to use the code.

In due course we hope to provide more complete documentation.

```idl
; Start IDL selecting both HELIO and EGSO
setav helio egso

;---------------------------
; The first object allows you to look at light curves and event lists to establish
; a time interval, select a set of instruments and retrieve the list of UMLs
hlio_setup = obj_new('hlio_setup')
hlio_setup -> set_base_time
hlio_setup -> set_base_time, timesrange=times

; ** Note: If you have not set a base time it will ask for one when you plot, etc.

hlio_setup -> plot
hlio_setup -> plot, /proton
hlio_setup -> plot, /both

hlio_setup -> select_times
; use cursors to select interval

hlio_setup -> load_eventlist
hlio_setup -> load_eventlist, name='goes_proton_events'
hlio_setup -> load_eventlist, /choose ['/me']);/f/ares],/particle

hlio_setup -> select_event
; select the instruments that you are interested in

hlio_setup -> select_instruments
hlio_setup -> select_instruments, /widget

; Retrieve list of observations that match time range and instrument selection

hlio_setup -> summary
; check on time range and instruments

filelist = hlio_setup -> get_filelist() ; output the list of UMLs
```
CASSIS is the Coordination Action for the integration of Solar System Infrastructures and Science
- CASSIS is funded under Research Infrastructures within the Capacities programme of FP7
- Started 1 June 2010, duration 36 months

CASSIS intended to facilitate science within the Solar System by improving the interoperability between data and services in all relevant domains

CASSIS brings together groups that are relevant to the issue
- HELIO, EuroPlanet RI and SOTERIA – all funded under FP7
- A number of key groups involved – ESA, NASA and NOAA
- Desire is to engage as many other groups as possible in the discussions, from Europe and the rest of the world

Web site: http://www.cassis-vo.eu

Solar System Science has traditionally been undertaken within a number of separate communities and disciplines

Increasing desire to do interdisciplinary science – science that spans the boundaries of the communities
- Many aspects of the system are inter-related
- Difficult to address them because of the lack of the integrating tools

Advances in technology means that some intrinsic differences between disciplines are being addressed
- Storage of and access to data should be simple mechanics
- Difficulties in access manifest by differing data formats and other dependencies

CASSIS encouraging interoperability by asking communities to:
- Try to follow simple rules when create data and metadata
  - Always include a minimum set of information in agreed formats
- Provide access via a more coherent set of protocols
- Identify ways of sharing services relevant to several communities
Cooperation & Discussion

- **Areas of cooperation include:**
  - Investigating ways to improve the interoperability between data and metadata from the domains, and the possibility of sharing services, including metadata resources.
  - Coordinating the use of standard within the projects and reflect any changes that are required to organizations like the IVOA and IPDA.
  - Coordination of the dissemination activities of the projects in order to create a more coherent and comprehensive approach.

- **Two principle means of discussion:**
  - *Community Consultation Meeting* will be used to gather input from the wider community.
  - *Vision for Solar System Science Workshops* are planned to bring key players together in order to lobby the case for solar system science with the decision makers and funding agencies.

Conclusions

- **HELIO development** is at the stage where it is becoming a useful tool for the heliophysics community.
  - Need user input to help us refine the capabilities.

- **Starting a series of Coordinated Data Analysis Workshops (CDAWs)** that will allow us to demonstrate HELIO’s capabilities and allow users to address science problems:
  - First one held at Trinity College Dublin (April 11-13)
  - Further CDAWs planned for autumn 2011 and spring 2012
  - Find out more in the Community section of the HELIO Web pages

- **CASSIS** is defining a better way of storing and accessing data.
  - Find out more on the CASSIS Web site
Work Packages

- **WP1 – Project Management**
  - General management of the project

- **WP2 – Interoperability of Data and Services**
  - Looking at ways of improving the quality and contents of metadata and data in order to improve interoperability
  - Examining ways of sharing services

- **WP3 – Networking**
  - Talking to the community to determine user requirements, etc.
    - Community should be varied and as international as possible

- **WP4 – Dissemination**
  - Producing various outreach activities at the international level
    - Web Portal, European Solar System Media Centre
    - Vision for Solar System Science Workshops
      - Invite key players and the national, European and international level
### Participants

<table>
<thead>
<tr>
<th>Short Name</th>
<th>Organization</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCL</td>
<td>University College London (MSSL and Physics &amp; Astronomy)</td>
<td>UK</td>
</tr>
<tr>
<td>KULeuven</td>
<td>Katholieke Universiteit Leuven</td>
<td>BE</td>
</tr>
<tr>
<td>FHNW</td>
<td>Fachhochschule Nordwestschweiz</td>
<td>CH</td>
</tr>
<tr>
<td>ROB</td>
<td>Koninklijke Sterrenwacht van Belgie</td>
<td>BE</td>
</tr>
<tr>
<td>INAF</td>
<td>Istituto Nazionale di Astrofisica (IFSI and Obs. Trieste)</td>
<td>IT</td>
</tr>
<tr>
<td>UPST</td>
<td>Université Paul Sabatier Toulouse III (CESR)</td>
<td>FR</td>
</tr>
<tr>
<td>OBSParis</td>
<td>Observatoire de Paris (LESIA)</td>
<td>FR</td>
</tr>
<tr>
<td>UCLA-IGPP</td>
<td>University of California, Los Angeles (IGPP)</td>
<td>US</td>
</tr>
<tr>
<td>RPI</td>
<td>Rensselaer Polytechnic Institute</td>
<td>US</td>
</tr>
<tr>
<td>ESA</td>
<td>European Space Astronomy Centre</td>
<td>ES</td>
</tr>
<tr>
<td>NOAA</td>
<td>Space Weather Prediction Center</td>
<td>US</td>
</tr>
<tr>
<td>NASA</td>
<td>GSFC Heliophysics Science Division</td>
<td>US</td>
</tr>
</tbody>
</table>

Participants include groups from the three project – HELIO, Europlanet RI and SOTERIA. Groups from the US were added during the negotiation phase. CASSIS interested in broadening the international aspects of the project by extending invitations to key organisations elsewhere. Whether other disciplines should be included is being considered – e.g. terrestrial climate studies.

---

**Things to look out for**

- Are the services behaving as you expect
- Are there and values that are obviously wrong
- ...
Instrument Capabilities Service (ICS)

- The ICS contains information about all the instruments that could be included in HELIO and is used to help determine which could have made the desired type of observations.

- The ICS knows characteristics of the observatories
  - When the observatory was operational
  - Observatory type and general location
    - Ground-based, Earth’s environment, heliosphere, planetary environment

- The ICS knows characteristics of the instruments
  - Observatory that the instrument is associated with
  - When the instrument was operational
  - Type of observation the instrument makes (remote/in-situ)
  - What the instrument is trying to observe
  - Type of instrument, wavelength etc. range, etc.

- Provider Access Table shows which can be accessed by HELIO

Instrument Location Service (ILS)

- The ILS is designed to help determine if an observatory was in the right place at the right time to observe the phenomenon.

- The ILS is still being developed; currently it knows the locations of the planets and interplanetary probes.

- Eventually the ILS will be able to also:
  - Determine whether the observatory lies within a region-of-interest
    - Location of the propagating phenomena determined by modelling
  - Calculate the location of satellites in the orbit around a planet
    - Location with respect to magnetospheric or ionospheric boundaries needs more information
Data Provider Access Service (DPAS)

- HELIO can access data from many domains using a variety of different access protocols
- Which instruments are available through HELIO and how they are accessed is defined by the Provider Access Table (PAT)
- Concentrated on adding larger providers to maximize return:
  - VSO, CDAWeb, PDS and PSA, BASS2000
  - Few smaller archives also available
- There are issues with each of these and future development of the DPAS will try to resolve these:
  - Planetary data are through PDS & PSA; URLs of files are stored in UOC because of complexity of access
  - Need to make changes to the PAT and ICS to accommodate structure of the CDAWeb; some parts hard to understand...
  - VSO good for most things; problems for some data and does not provide required access in some cases
  - BASS2000 would be much easier with small change to archive...

HEC, MDES and HFC

- Heliophysics Event Catalogue (HEC)
  - The HEC hosts event catalogues of different types from many sources and communities
    - Events detection based on some criteria (which may be wrong!)
    - Data are conditioned as they are ingested
    - Process of defining additional annotation is underway
- Metadata Evaluation Service (MDES)
  - The MDES allows users to search time-series data for events that are subtle and may need some type of processing to aid detection
    - Recognition that not all events can be detected a priori
  - Procedures used to detect events can be stored and reused
  - Outputs are event lists in their own rights (stored as VOTables)
- The Heliophysics Feature Catalogue (HFC) may also be of use
  - Characterises features and their motion on Sun and in heliosphere