The Heliophysics Integrated Observatory HELIO

Robert Bentley, UCL-MSSL (UK)

HELIO Project Coordinator

http://www.helio-vo.eu

(IVOA, 10 November 2009)

Overview

- Heliophysics is the study of the effect of the Sun on the Solar System
  - New discipline that spans a number of well established communities
- HELIO, the Heliophysics Integrated Observatory, 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 a capability of this type is driven by desire to study problems that span disciplinary boundaries
  - Search base on metadata increasingly important as data volumes increase
- HELIO 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
- Strong desire is to create something in useful to the community
  - Community input and feedback needed during all phases of the project

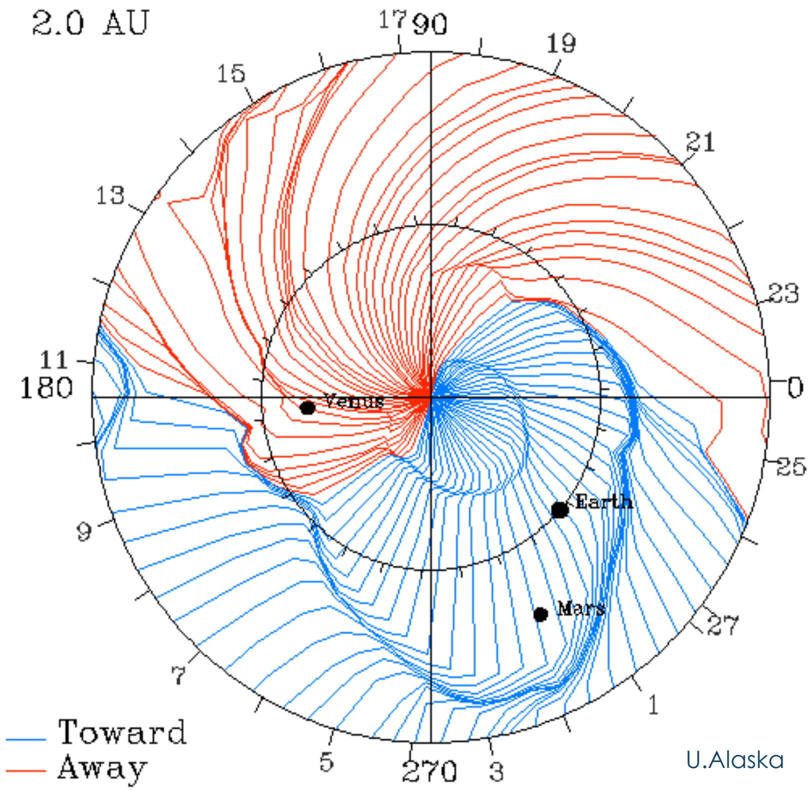
Objectives and Structure

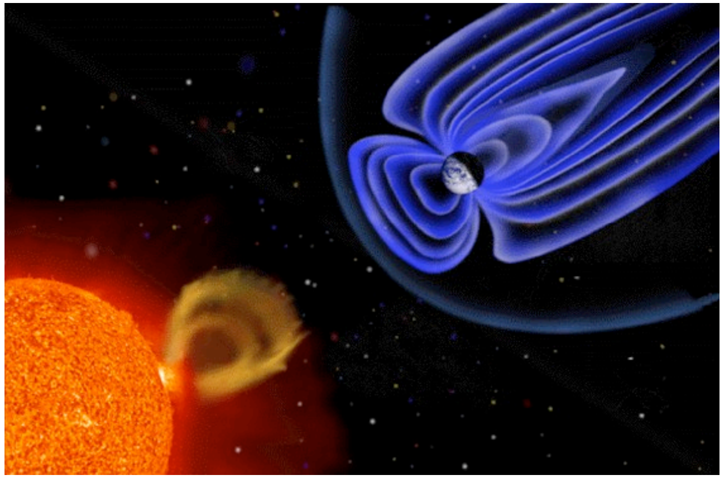
- **Project Objectives**
  - To create a collaborative environment where scientists can discover, understand and model the connection between solar phenomena, interplanetary disturbances and their effects on the planets (esp. the Earth)
  - To establish a consensus on standards for describing all heliophysical data and champion them within international standards bodies, e.g. the IVOA
  - *To develop new ways to interact with a virtual observatory that are more closely aligned with the way researchers wish to use the data*

- **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

- *In the rest of the talk we will examine:*
  - *The context in which HELIO operates*
  - *How the three Activities are used to address the problems*

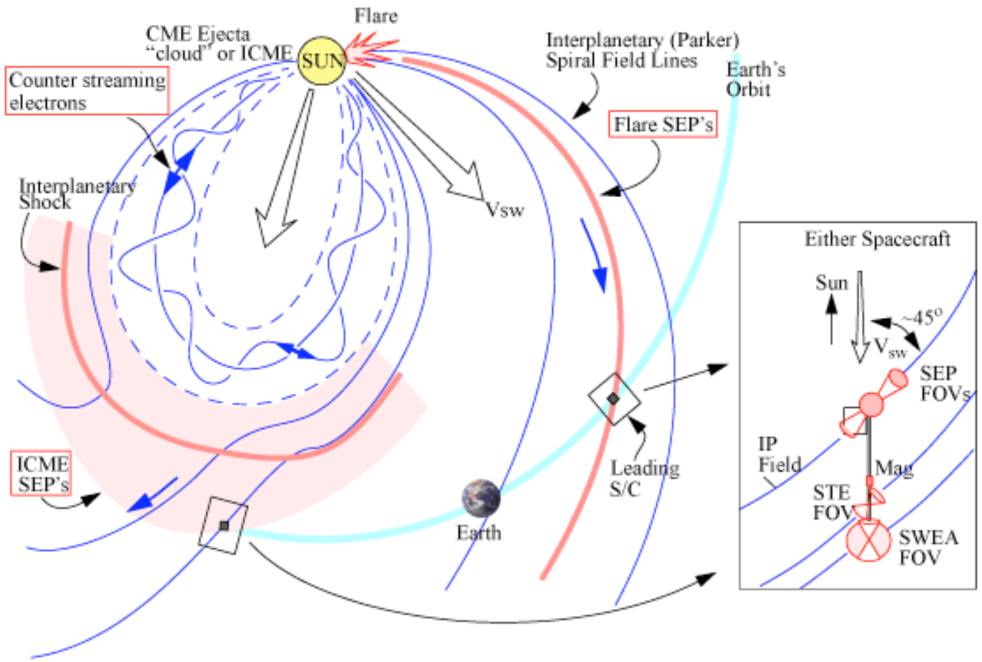
Context and Communities





- **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 effects related to propagating phenomena resulting from activity
  - Immediate and delayed effects result from the different types of emission
- **Location of observer in relation to the source and with respect to a planet determines nature of effects**
  - Presence of magnetic field and/or atmosphere influences effect on planetary environment

Context and Communities



- Immediate effects relate to photons emissions
  - Short delay due to light propagation time
  - Effects mostly experienced if observer has line-of-site view of the source
- Delayed effects cause by particle emissions
  - Particles follow the spiral of the IMF – effects lags behind source
  - Coronal Mass Ejections (CME) distort IMF and carry embedded field
  - Delay can be tens of hours, even at Earth
  - Phenomena must pass observer for effect to be experience

- Different aspects of effects traditionally studied in separate domains
  - Solar, heliospheric, magnetospheric & ionospheric (planetary & geo-sciences)
  - Communities have evolved independently over decades or centuries
  - Significant differences in the content of data, and the way they are stored and handled

HELIO's perspective of a generic problem

- **Identify interesting things to study**
  - Event lists and feature lists used as primary selection criteria
  - Search undertaken across several domains in 4-Dimensions
    - Effects occur as phenomena propagate – whether, where and when to look
    - Follow phenomena through coordinate systems as enter planetary environment
  - *Search based solely on metadata and derived products*
- **Review availability of suitable observations**
  - Determine whether observations matching selection criteria are available for the relevant locations and time intervals
  - 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)
  - Optionally process selected observations (extract and calibrate)
  - Optionally return data in different/desired format
- ***Analysis done with users own software tools (e.g. IDL)***

Implementation – Service Activities

- **HELIO is based on a Service-Oriented Architecture**
  - **Series of search metadata services covering the different domains**
    - All types of Event and Feature metadata available
    - Context information helps user to refine selection – time series, images, etc.
  - **Services to identify and retrieve observations based on search results**
    - Determines whether suitable observations were made at particular time/place
    - Knows which data are stored where and how to access them
  - **Enabling services provide processing and storage capabilities**
    - Tools provided to find and track events/phenomena in 4D environment
    - Tools to generate quick-look products, etc.
    - Tools to extract & calibrate subset of data in any desired format
  - **Service to provide enhanced search capability based on metadata**
    - HELIO GUI sits on top of the API to this service
    - Facilitates search across domains for interesting events
    - Propagation model available to relate metadata from all part of Solar System
  - ***Services can be used individually or combined through workflow capability***
  - **Semantic-driven approach used to integrate data from different domains**
    - Based on ontology derived from existing data models

Services that support the search

Search Metadata	
Heliophysics Event Catalogue (HEC) Service	Maintain and provide access to *existing* event data from all domains
Heliophysics Feature Catalogue (HFC) Service	Maintain and provide access to *existing* feature data from all domains
MetaData Evaluation Service (MDES)	Allows the user to create an auxiliary event list based on a *newly-derived* parameter, etc.
Context Service (CXS)	Provide context information to help the user make a selection
Ancillary Information Service (AIS)	Provide integrated access to external resources that do not conform to HELIO interface standards

- **The HEC and HFC provide criteria for making the selection of a time interval**
- **The MDES allows the user to generate events that are based on new criteria or data (cf. AMDA); it also facilitates the joining of metadata lists**
- **The AIS affords access to capabilities outside of HELIO (e.g. SPIDR & AMDA) by wrapping them and making them *HELIO compliant***
- **Context information (CXS) helps the user refine their selection**

Services to select observations

Review suitable observations		
	Instrument Capabilities Service (ICS)	Match required observation type to one or more instruments (each part of an observatory)
	Instrument Location Service (ILS)	Determine the location of an instrument (part of an observatory) at a specified time
	Observation Coverage Service (OCS)	Provide information on whether an instrument was making suitable observations at a specified time
Locate and Retrieve the Data		
	Data Provider Access Service (DPAS)	Provide integrated access to data archives in all domains no matter how the data are stored or accessed

- **Combination of the ICS and ILS allows user to identify which instruments that make the type of observations that are required were located where they could have made suitable observations**
- **The OCS determines whether they were observing during the specified time interval (at some granularity)**
- **Selected observations can then be retrieved by the DPAS**
  - **Knows where and how data are stored around the globe**
  - **Has a Data Registry to keep track of data resources**

Enabling Services

Enabling Services	
Processing Service	Support processing on demand
Storage Service	Provide storage for user information
Coordinate Transformation Service (CTS)	Translated between the different coordinate systems used by the communities
Semantic Mapping Service (SMS)	Maps terms used in the metadata from the different communities
HELIO Registry Service (HRS)	Maintain and provide access to a registry that describes all the services available to HELIO
HELIO Monitoring Service (HMS)	Keeps track of the status and performance of the services that the HRS knows about
Community Interaction Service (CIS)	Manages interactions with the community, including authentication and usage statistics

- **These are low level services that underpin the other services**
- **The Registry Service (HRS) is used to allocate services for the HELIO Search Engine (or any other system that wishes to employ HELIO services)**
- **The capabilities of the CTS and SMS could be need in various places and may be implemented as sets of library routines**

The HELIO Search Engine

Top-level Utility		
	HELIO Search Engine	Provides an integrated advanced search capability;

- **The HELIO Search Engine (HSE) integrates the services using the *Taverna* workflow tool**
  - **Which services are used and the order that the deployed in can be altered within the workflow by the user. HELIO will provide a number of pre-canned workflows that will satisfy the need of many users**
  - **A propagation model relates observations made in different parts of the solar system**
  - **The main HELIO API and GUI are interfaced with/controlled by the workflow tool and provide the main means of access for the user**

- **A user could also employ their own workflow tool or scripting capability to combine the services**
  - **Services are allocated in response to requests submitted to the Registry Service (HRS)**

- **General notes about implementation:**
  - **Each service has its own API and can be called individually**
  - **Services will often be co-located to improve the efficiency of the system**
  - **Multiple instances of each service will improve system resilience**

Use Case

Study flares observed in hard X-rays (RHESSI) and in radio waves (GBO XX)

- **Heliophysics Event Catalogue (HEC) service**
  - **Find flares (type?) in specified time range (tstart, tend)**
    - **What qualifiers (threshold, location, etc.)?**
- **MetaData Evaluation Service (MDES)**
  - **Filter HEC results with daily observing period defined by the ILS/ICS**
  - **Filter for both instruments observing as defined by the OCS**

⇒ List of jointly observed flares

- **Context Service (CXS)**
  - **Helps user decide which events may be of most interest**
- **Data Provider Access Service (DPAS)**
  - **Retrieve selected data sets (optionally process)**

Implementation – Research Activities

Research Activities used to investigate/develop required capabilities

- **Develop ontology to create over-spanning data model (Semantic VO)**
  - Based on existing data models such as EGSO, VSTO and SPASE
  - Not practical to develop single data model for everything, nor is it realistic to expect communities to adopt it over their own models
- **Feature detection and tracking capabilities in new domains**
  - Features in 2D images, time-spectra data, time-intensity data…
- **Enhanced user interface to encourage uptake of VO capabilities**
  - We believe that we need to find ways of making the interfaces of Virtual Observatories match the way scientists wish to think…

Implementation – Networking Activities

- **Involving the Community**
  - **Community consulted to determine project requirements**
    - Focus Groups used to resolve issues and formulate designs
    - User Groups used to review capabilities and provide feedback
  - **Coordinated Data Analysis Workshops (CDAWs) are planned**
    - Provisional dates are in months 18, 24 and 30
  - **Community Coordination Meetings will also be held**
    - Some will be in conjunction with meetings like the EGU, AGU, etc.

- **Strategy and Standards**
  - **HELIO will provide integrate access to metadata and data from several domains that are related but have distinct differences**
  - **Issues related to the quality and coherence of some metadata**
    - Semantic-driven approach will resolve many, but not all problems
  - **Need to improve the search and observational metadata**
    - Regenerate some metadata – event lists, etc.
    - Talk with data providers to ensure header information improved
    - Develop new standard for future datasets; accommodate existing ones
  - **Make improvement in collaboration with the global community**
    - Compliant as far as possible with standards develop by IVOA, etc.
    - Need to be sympathetic to related communities - astrophysics, planetary, etc.

The HELIO Consortium

Country	Institution	Contact
UK	University College London (MSSL)	R.D. Bentley (PI)
CH	Fachhochschule Nordwestschweiz	A. Csillaghy
FR	Observatoire de Paris (LESIA)	J. Aboudarham
FR	Universite Paul Sabatier Toulouse (CESR)	C. Jacquey
UK	Science and Technology Facilities Council (RAL)	M.A. Hapgood
FR	Universite Paris-Sud (IAS)	K. Bocchialini
IT	Istituto Nazionale di Astrophisica (Obs. Trieste)	M. Messerotti
UK	University of Manchester	J. Brooke
IE	Trinity College Dublin	P. Gallagher
US	Rensselaer Polytechnic Institute	P. Fox
US	Lockheed Martin Space Systems Company (LMATC)	N. Hurlburt
US	National Aeronautics and Space Administration *(Heliophysics Science Division at GSFC)*	D.A. Roberts
Int.	European Space Agency *(Science Operations Dept., Space Environment & Effects Dept.)*	L. Sanchez

- Several partners have extensive knowledge of the domains of heliophysics and experience in managing and archiving their data
- Others provide expertise in different areas of computer science and also have some familiarity with handling data

HELIO and the IVOA

- In implementing HELIO, considering various IVOA standards
  - VOEvent, VORegistry, VOSpace, workflow, etc, etc…
  - Already using VOTable extensively

- Intension is to use the standards where we can and report the required modifications back to the IVOA
  - Inevitably some dependencies on domains specific requirements
  - Differences because heliophysics mainly an event driven science

- HELIO will also try to coordinate the requirements of the communities that constitute heliophysics to the IVOA, etc.
  - Desire to to further generalize the IVOA standards
  - Need to be sensitive to the needs of other related communities
    - Heliophysics sits between astrophysics and geo- and planetary sciences