

Building Blocks for a Virtual Observatory in Heliophysics

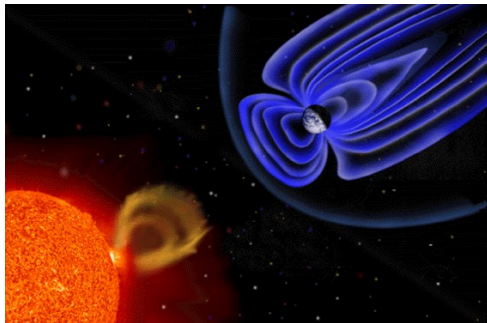
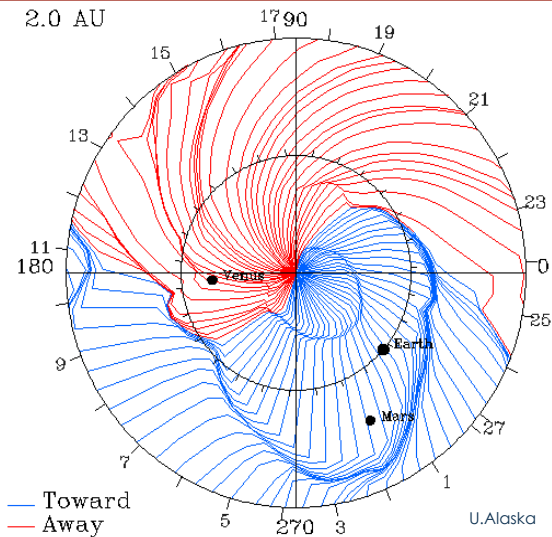
Robert Bentley

University College London

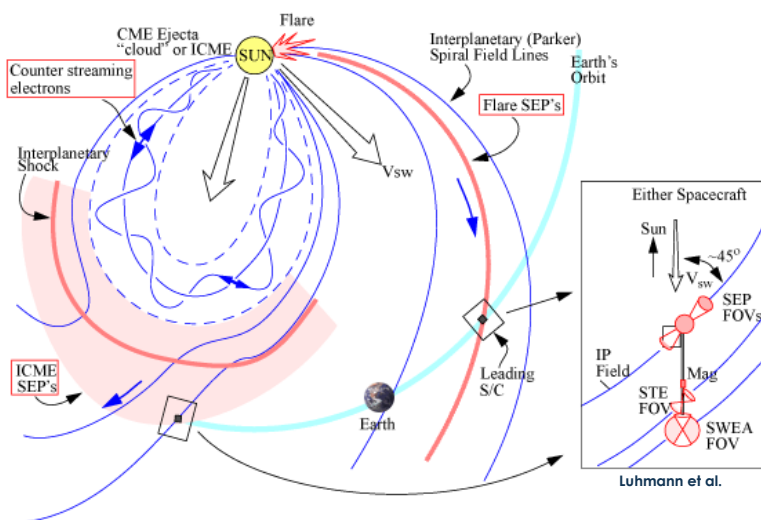
Coordinator of HELIO and CASSIS FP7 projects

(Fall AGU, San Francisco, 9 December 2011)

- **The Heliophysics Integrated Observatory, HELIO, is a Research Infrastructure funded under EC's FP7 Capacities Programme**
- **HELIO is creating a collaborative environment where scientists can discover, understand, and model the connection between solar phenomena, interplanetary disturbances, and their effects on planets, especially the Earth**
 - **Need for capability driven by desire to study problems that span disciplines**
 - **Search base on metadata increasingly important as data volumes increase**
- **HELIO provides the ability to identify interesting phenomena and access relevant observations**
 - **solar and heliospheric, together with related geophysical and planetary data (magnetospheric and ionospheric)**



- **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...**



Location is important

- **Effects depend on type of emission**
- **Propagation is influenced by the interplanetary magnetic field for particles**
- **Modelling needed to help understand when, where and whether to look**

- **Identify interesting things to study**
 - Science Objectives define what user looking for
 - Search undertaken in 4-Dimensions across several domains
 - Effects occur as phenomena propagate and interact
 - Modelling required to understand **whether, where and when** to look
 - 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
 - Determine whether the instruments were making observations
 - Decide if they could be showing something of interest

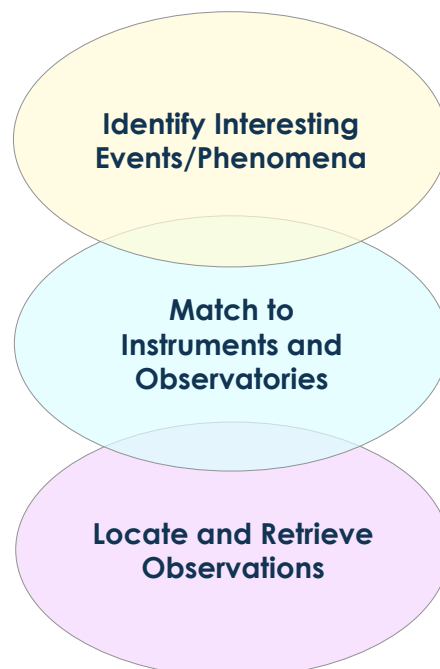
- **Locate, select and retrieve the required observations**
 - For **all domains**, system knows which types of data are held where and handles access no matter how data are stored (access protocols & formats)

- **Analysis done with users own software tools (e.g. IDL/SSW)**

HELIO is an **integrated system** implemented with a **service-oriented architecture**

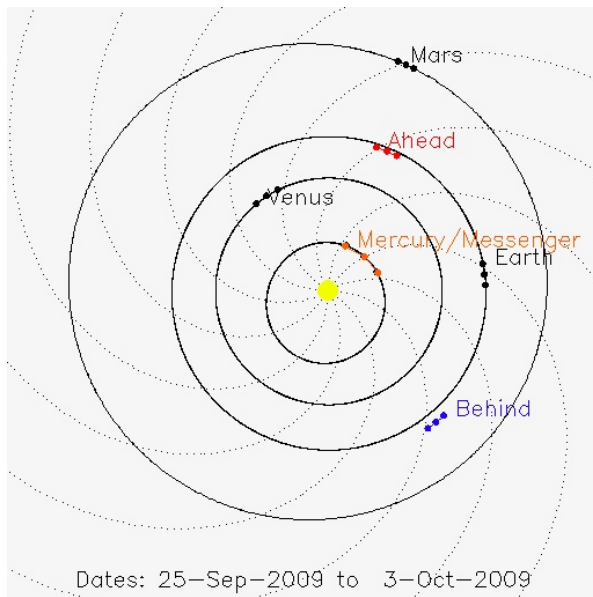
Design splits the tasks into a set of components or **services**

- Services to aid search process and 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



- **Objective:**
Aspects of phenomenon → Set of Obs. Types @ Times, Locations
- **HELIO provides as diverse a set of tools as possible, based on metadata from all relevant domains**
 - **Heliophysics Event Catalogue (HEC)**
 - Catalogue of existing event data
 - **Data Evaluation Service (DES)**
 - Allow user to evaluate time series data and generate own event list
 - **Heliophysics Feature Catalogue (HFC)**
 - **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**

- **Objective:**
Obs. Type @ Time, Location → Instrument/Observatory @ Time
- **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 and hence the instruments
 - **Observation Coverage Service (OCS)**
 - **Unified Observing Catalogue** used to handle special cases
 - **Observation Coverage Table** determines if observatory active
- **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**



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

- User identifies an event in one part of the solar system (HEC) 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. UOC needed for planetary data

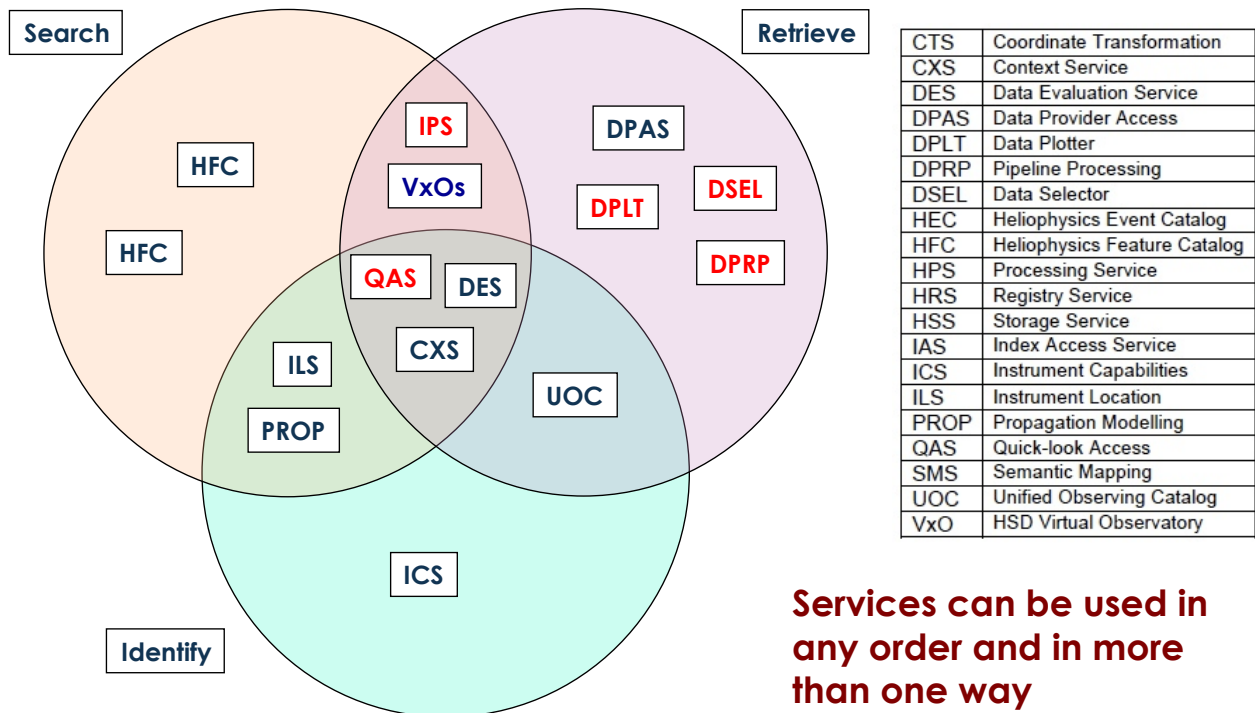
Defined Services

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, coordinate transformation, security, etc.

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

Service Name	Purpose
<i>Search Metadata</i>	
Heliophysics Event Catalogue (HEC)	Maintain and provide access to <i>existing</i> event data from all domains
Heliophysics Feature Catalogue (HFC)	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 <i>newly-derived</i> parameter, etc.
Context Service (CXs)	Provide context information to help the user make a selection
<i>Review suitable observations</i>	
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
<i>Locate and Retrieve the Data</i>	
Data Provider Access Service (DPAS)	Provide integrated access to data archives in all domains no matter how the data are stored or accessed
<i>Enabling Services</i> ←	
HELIO Registry Service (HRS)	Maintain and provide access to a registry that describes all the services available to HELIO
Community Interaction Service (CIS)	Manages interactions with the community, including authentication and usage statistics
Processing Service	Support processing on demand
Storage Service	Provide storage for user information
Ancillary Information Service (AIS)	Provide integrated access to external resources that do not conform to HELIO interface standards
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 Monitoring Service (HMS)	Keeps track of the status and performance of the services that the HRS knows about
Resource Usage Service (RUS)	Keeps track of usage of HELIO so that the project can provide statistics to users, providers, etc.



The Virtual Observatories have been developed as an alphabet soup supporting part of the heliophysics...

- **VxOs etc. under NASA-GSFC Heliophysics Science Division:**
 - Virtual Solar Observatory (VSO)
 - Virtual Heliospheric Observatory (VHO)
 - Virtual Space Physics Observatory (VSPO)
 - Virtual Magnetospheric Observatory (VMO)
 - Virtual Ionosphere Thermosphere Mesosphere Observatory (VITMO)
 - Virtual Radiation Belt Observatory (ViRBO)
 - Virtual Wave Observatory (VWO)
- **VxO project funded by NSF:**
 - Virtual Solar Terrestrial Observatory (VSTO)
- **Other relevant (resource rich) data providers**
 - National Space Science Data Center (NSSDC)
 - NASA's Space Physics Data Facility (SPDF)
 - NOAA's National Geophysical Data Center (NGDC)

- The capabilities that HELIO is developing should be thought of a building block in a larger capability – parts of a tool kit...
- **Service-oriented architecture has advantages**
 - Services can be used individually or as part of a workflow
 - Method of implementation is hidden from the user
 - New capability can be implemented as a new service
 - Services can be developed and maintained independently
- **Services interfaces need to be compliant with a set of standards in order to ensure interoperability**
- **HELIO has defined Web Service interfaces for its services based on IVOA specifications (with extensions)**
 - If adopted by other the VxOs and other capabilities, these could become part of the tool kit
 - *May need some iteration to satisfy needs of all*

- **HELIO Query Interface (HQI) used on most services**
 - Input based on the Parameter Query Language (PQL) and Table Access Protocol (TAP)
 - Output uses **annotated** VOTable
 - Both synchronous and asynchronous requests – latter known as “long running query”
- **Services that involve processing follow the Universal Worker Service (UWS) pattern**
 - Input usually via an XML file (more flexible)
 - Output could be data file (VOTable or otherwise), or an image
 - Wait for completion flag from service
 - Services include Context Service and Coordinate Transformation
- **Services usually have both SOAP and REST interfaces**
 - REST interface used to implement Test GUI for service

- **Quality of available metadata is very variable**
 - Dependant on the community involved
 - Some projects are better than others
- **As metadata has been ingested into the HELIO services**
 - In HEC time and coordinates sometimes needed attention
 - Names used for parameters were standardized
 - In UOC tried to standardize way observations described
 - Times, fields of view, etc.
- **Problem partly because of lack of standards or guidelines**
 - There are standards defined by a number of organizations
 - No overall agreement on what to use!!!!
- **Urgently need to address the problem**

- **Increased interoperability needed between the capabilities developed by the heliophysics community**
 - Existing capabilities start to become part of a larger picture
 - Framework in which to contribute new capabilities (cf. SolarSoft)
 - **Everyone would benefit !!!!!**
- **Making the services and other capabilities interoperable is only the start**
 - Need to improve the quality and contents of the **metadata**
 - Needs to improve the quality of the **data**
- **The heliophysics community need to act on this sooner rather than later**
 - New capabilities should be part of the larger picture

- **HELIO has been developed using a service-oriented architecture that could be used as the base on which to establish a framework for future development**
- **The heliophysics community would gain significantly if it tried to work together towards this end**
 - Currently our efforts are too fragmented
 - Ability to do science is being constrained
- **Adoption of standards and improved interoperability desirable no matter what we decide to do**
 - Improved data and metadata quality will always be beneficial
- **Need to discuss how to proceed**
 - Time for a meeting???