



# The Role of Virtual Observatories in Space Weather and HELIO Use Cases

**Robert D. Bentley**  
University College London

(ICTP, 26 October 2010)



*Heliophysics Integrated Observatory (HELIO)*



# Outline



- **Overview of the types of data providers**
  - Review of some European VOs
- **Description of the Heliophysics Integrated Observatory**
- **Standards and interoperability**
  - Issues related to use of data by HELIO
  - Description of CASSIS
- **Use Cases**
- **Summary**





## Many types of Data Provider



- **Data providers have different capabilities and resources**
  - User must interact with them in many different ways
    - Many transfer protocols – http, ftp, cgi-bin, Java-script, Web-service
  - Each serves a purpose and all are important
- **Resource-poor Providers**
  - Able to do little more than make the data available via the Internet
  - User has to do a significant amount of the work
  - **R-P Providers are still be a critical part of the overall picture!**
    - Can fill in the gaps in coverage for ground-based observations
- **Resource-rich Providers**
  - Capable, intelligent providers with some added value
    - Includes data centres and research infrastructures
  - Supply required data in response to a simple query
    - Some issues if interface, etc. not properly designed





## Many types of Data Provider



### Virtual Observatories (VO)

- **What is a Virtual Observatory?**
  - A virtual observatory (VO) is a collection of interoperating data archives and software tools which utilize the Internet to form a scientific research environment in which astronomical research programs can be conducted (Wikipedia)
- **More specifically VOs**
  - Are high-end resource-rich “providers” that add a lot of value
  - Provide integrated access to data from all types of provider, including other VOs
    - The source of the observations is a secondary issue
    - VOs provide a standardized interface for the user
  - Provide tools that can be used to mine the data
    - Tools allow searches based on science use cases
  - VOs do NOT gather data into a central archive
    - Search based on metadata (which can be accumulated) and on derived products





# European Archives



- **In Europe, most observations made by ground-based solar observatories are curated by the institutions**
  - Large number of small observatories scattered over the continent
  - Can fill in the gaps in coverage for ground-based observations
- **Some ground-based observations concentrated in data centres**
  - BASS2000 (France) is a data centre for (mostly) French observatories
  - SOLARNET (Italy) is a data centre for Italian observatories
- **There are several data centres that concentrate on space-based observations**
  - MEDOC (France) has data from SOHO, TRACE, etc.
  - UKSSDC (UK) concentrates on space science observations
  - SIDC (Belgium) concentrates on space weather data
  - PSA (Spain) archive of European planetary data
  - HSDCE (Norway) hosts data from the Hinode mission
  - HEDC (Switzerland) hosts data from the RHESSI mission





# European Virtual Observatories



- **Several European research infrastructures – each covers a different domain and there are differences in emphasis and approach**
  - **HELIO – Heliophysics Integrated Observatory**
    - HELIO is building a virtual observatory for heliophysics
  - **SOTERIA – SOLar TERrestrial Investigations and Archives**
    - SOTERIA focuses on the Sun-Earth connection
  - **SWENET – Space Weather European NETwork**
    - SWENET focuses on data related to Space Weather
  - **Europlanet RI – Europlanet Research Infrastructure**
    - Europlanet concerns itself with bodies other the Earth
  - **Euro-VO – European Virtual Observatory**
    - Euro-VO deals with the astrophysical side of astronomy
  - **VAMDC – Virtual Atomic and Molecular Data Centre**
    - VAMDC handles need for atomic and molecular data
- **CASSIS – Coordination Action for the integration of Solar System Infrastructures and Science**
  - CASSIS is a discussion forum that deals with the more general case
  - Examines issues related standards and interoperability
  - Grown out of cooperation between HELIO, SOTERIA and Europlanet RI





# Data Centres in the US



- **Many relevant data providers**
  - STEREO Science Center (NASA)
  - RHESSI Data and Software Center (NASA)
  - SDO Joint Science Operations Center (LMATC and Stanford)
  - Etc...
  
  - SDAC – Solar Data Analysis Center (NASA)
  - SPDF – Space Physics Data Facility (NASA)
  - PDS – Planetary Data System (NASA; many nodes)
  
  - National Space Science Data Center (NSSDC)
    - Permanent archive for NASA's space science mission data.
  - National Geophysical Data Center (NGDC)
    - Permanent archive for NOAA's geophysical data collection





# Virtual Observatories in the US



The Virtual Observatories have been developed within NASA, etc. as an alphabet soup...

- **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)
- **The main problem is how to use them together...**





- **The Space Weather European Network (SWENet) has been operational since 2005 and was established under the programme of ESA's Space Weather Applications Pilot Projects.**
  - Central resource for European space weather activities
  - Provides access to space weather data and services
- **SWENET provides the ability to**
  - View the current space weather situation
  - Search for data and services
  - Browse the SWENET database with customized retrieval capabilities
- **The SWENET portal provides access to produce and services developed by Service Development Activities (SDAs)**
  - The SDAs focus on a wide range of space weather domains that to address the needs of specific users

- EuroPlaNet was a 4 year Coordination Action funded under FP6
- The Europlanet Research Infrastructure (RI) is a follow-on project funded under the Capacities specific programme of FP7
  - Started January 2009, duration 48 months
- Europlanet RI follows the Integrated Infrastructure Initiative (I3) activities model, but with a different emphasis to HELIO:
  - Networking Activities are aimed at further fostering a culture of cooperation in the field of Planetary Sciences
  - Trans-national Access Activities are to provide:
    - Trans-national access to a range of laboratory and field site facilities tailored to the needs of planetary research
    - On-line access to the available planetary science data, information and software tools, through the IDIS e-Service
  - Joint Research Activities are aimed at improving the services provided by the ensemble of Trans-national Access Activities

- **The Virtual Atomic and Molecular Data Centre (VAMDC) is a Research Infrastructure funded under the Capacities specific programme of FP7**
  - Started July 2009, duration 42 months
- **VAMDC aims to build an interoperable e-Infrastructure for the exchange of atomic and molecular data**
- **VAMDC is funded under the same call as HELIO and also follows the Integrated Infrastructure Initiative (I3) activities model, but with a different emphasis:**
  - **Networking Activities** will coordinate the infrastructure activities among all trans-disciplinary fields both within the ERA and
  - **Service Activities** will create a unique, state-of-the-art e-infrastructure for both atomic and molecular data producers and users
  - **Joint Research Activities** will build the complete set of “tools” necessary to create the VAMDC e-Science platform



# Role of VOs in Space Weather



- **Space Weather (SWx) involves the study of the effect of solar activity on the Earth's environment**
  - **The effect on technology is an increasing concern**
- **We need to examine what is happening on the Sun to understand the phenomena that result in SWx effects**
- **We need to be able to monitor the Earth's environment to understand how it is reacting to external stimuli**
- **There is therefore a requirement to search through various sets of data looking for combinations of circumstances**
  - **To be able to really understand an effect we must identify many instances of its occurrence**
- **The virtual observatory paradigm is designed to facilitate this**
- **In Europe, two infrastructure are relevant – HELIO and SOTERIA**





# SOTERIA



- **The SOLar TERrestrial Investigations and Archives (SOTERIA) is a Space Science project funded under the SPACE call of FP7**
  - Started November 2008, duration 36 months
- **SOTERIA aims to improve our understanding of the space weather phenomena through collaboration between experts in different fields of solar, space, and geophysics**
  - Main goal is to provide databases that go beyond the state-of-the-art
    - Improved details, time-resolution and access methods
  - Considerable effort in utilizing the existing and developing improved theoretical and simulation models for interpreting the space weather data
- **Studies conducted by SOTERIA involve the:**
  - Analysis and processing of the relevant data from 18 satellites, including several ESA and other European satellites
  - Complemented by a large set of data from European ground-based observatories
- **SOTERIA is using database technology developed by HELIO**
  - Catalogues based on the Heliophysics Event Catalogue (HEC)





# HELIO



- **HELIO, the Heliophysics Integrated Observatory, is a Research Infrastructure funded under Capacities Specific Programme of the EC's Seventh Framework Programme (FP7)**
  - 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 elsewhere**
  - HELIO will create a collaborative environment where scientists can discover, understand and model the connection between solar phenomena, interplanetary disturbances and their effects on the planets
    - Ability to identify interesting phenomena and access relevant solar and heliospheric data and related data from the planetary environments
  - Need for the capability is driven by the desire to study problems that span disciplinary boundaries
  - A search base on metadata becoming increasingly important as the volumes of data increase



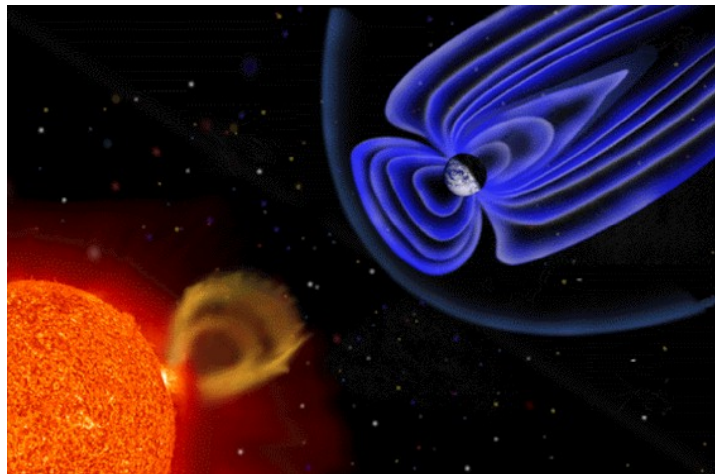
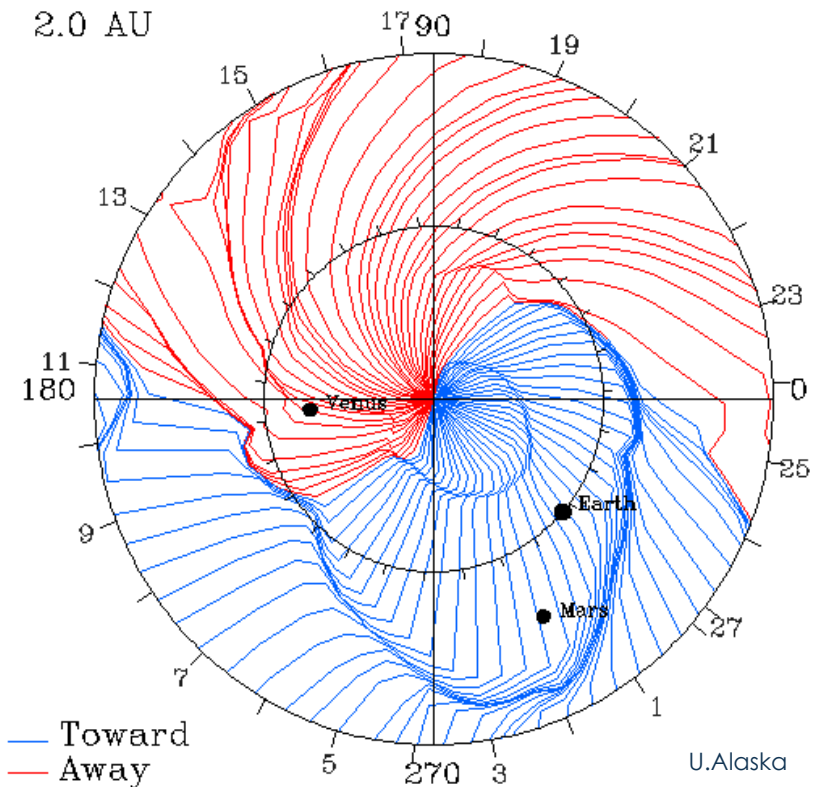


# HELIO



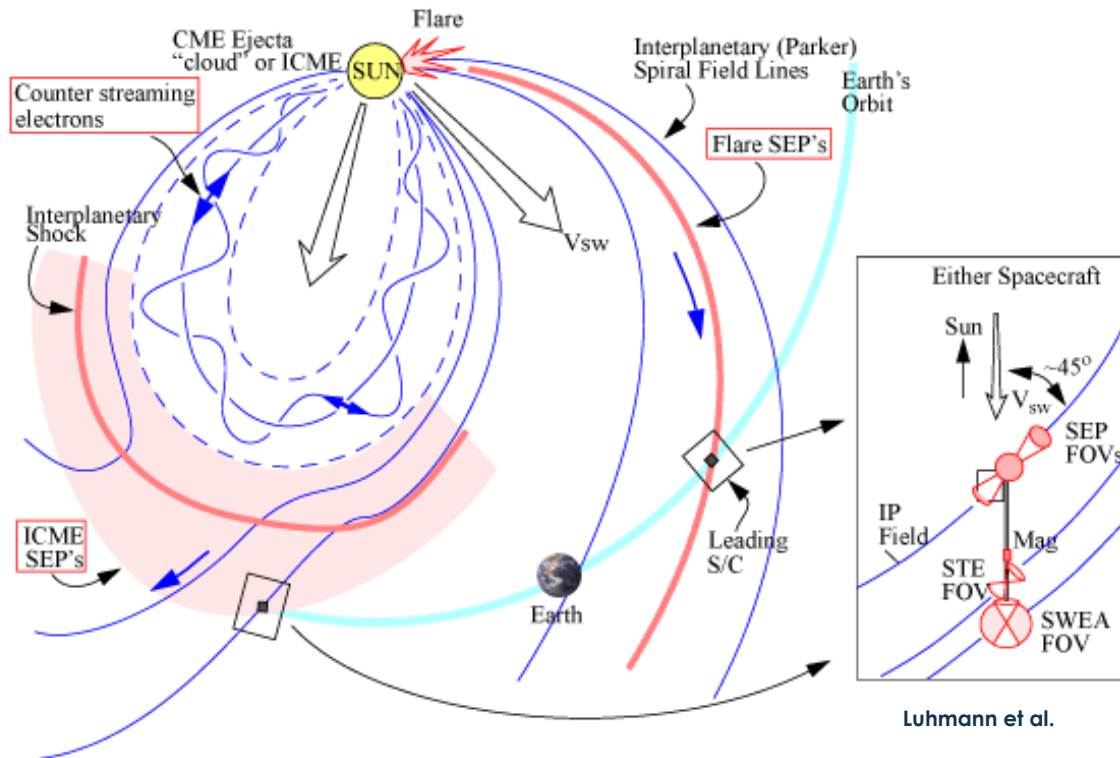
- HELIO will address its challenges following the **Integrated Infrastructure Initiative (I3) activities model of FP7**:
  - **Networking Activities** used to involve the community
  - **Service Activities** used to implement the infrastructure of the VO
  - **Research Activities** used to investigate/develop capabilities
- In order to explain how HELIO works we will concentrate mainly on the **Service Activities**





- **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
  - 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 what is observed
- **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...**





- **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 due to particle emissions**
  - Particles follow the spiral of the IMF – effects lags behind source
  - Coronal Mass Ejections (CME) distort the IMF and carry embedded field
  - Delay can be tens of hours, even at Earth
  - Phenomena must pass observer for effect to be experience

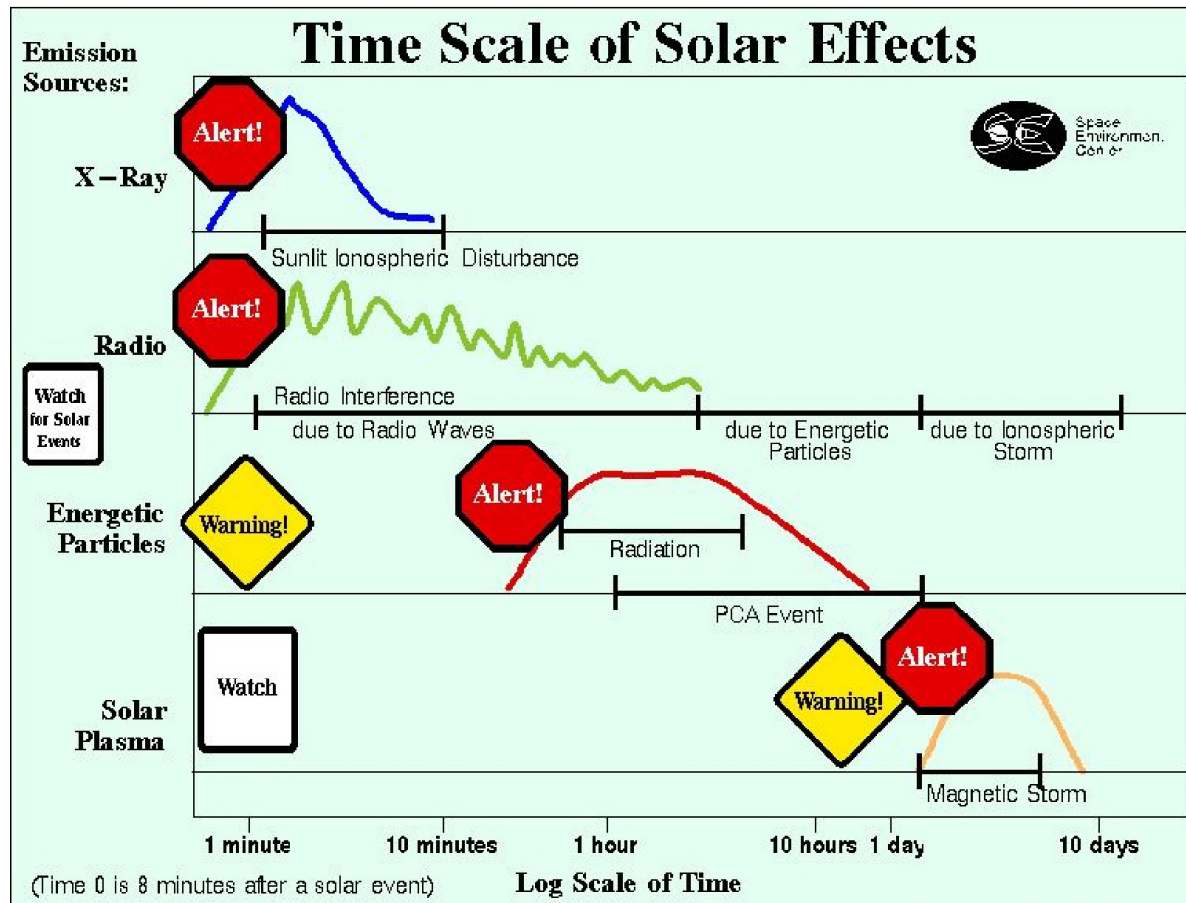


Figure produced by NOAA's Space Weather Prediction Center (SWPC; was SEC)

- The different types of emission produce the family of space weather phenomena observed at the Earth
- Similar phenomena are observed in other parts of the Solar system



# HELIO's perspective of a Generic Problem

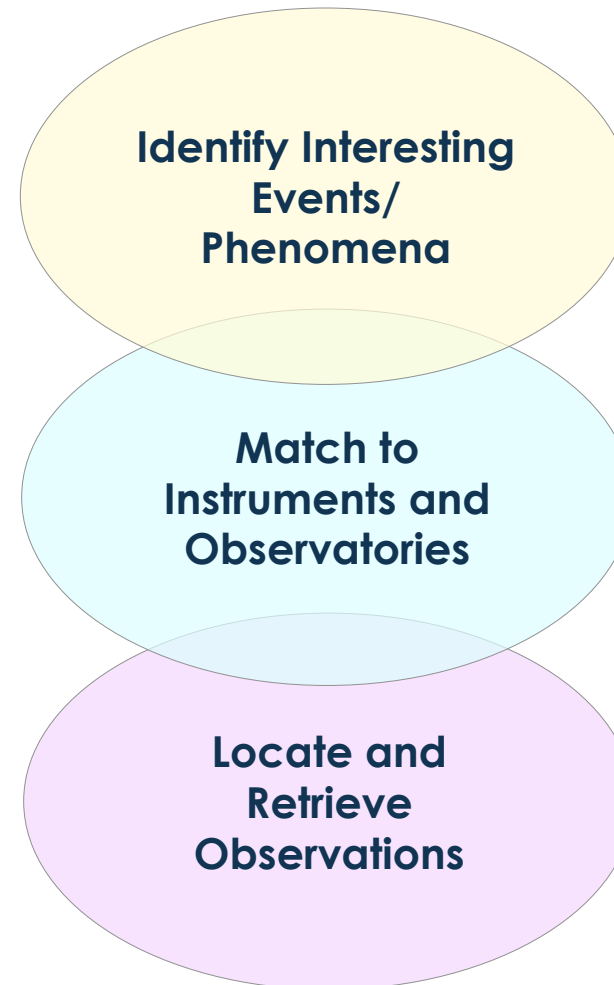


- **Identify interesting things to study**
  - Search undertaken in 4-Dimensions across several domains
    - Effects occur as phenomena propagate – whether, where and when to look
    - Follow phenomena through coordinate systems as they evolve
  - Search based solely on metadata and derived products
    - Event lists and feature lists from many domains used as a primary selection criteria
- **Review availability of suitable observations**
  - Determine whether suitable instruments at the relevant locations
    - Science objectives dictate types of observations required
  - Determine whether instrument was 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)
  - Optionally process selected observations (extract and calibrate)
  - Optionally return data in different/desired format
- **Analysis done with users own software tools (e.g. IDL)**



**Design of HELIO intended to split the task into components that can be developed and used independently or as part of a workflow**

- **Search process intended to refine selection of what is relevant to scientific objective**
  - Search based on metadata and derived products
- **Only locate and retrieve data when selection process complete**
  - Interaction with data providers minimized and very targeted
- **Tools to return only the required parts of the data**





# Search Process



- Objective is to turn a science use case that is looking at aspects of a phenomenon into sets of required observations at specific locations and times
- HELIO tries to provide as diverse a set of tools as possible, including 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 process



- 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





# SO Architecture

HELIO being implemented as a set of services service:

- Identify interesting things

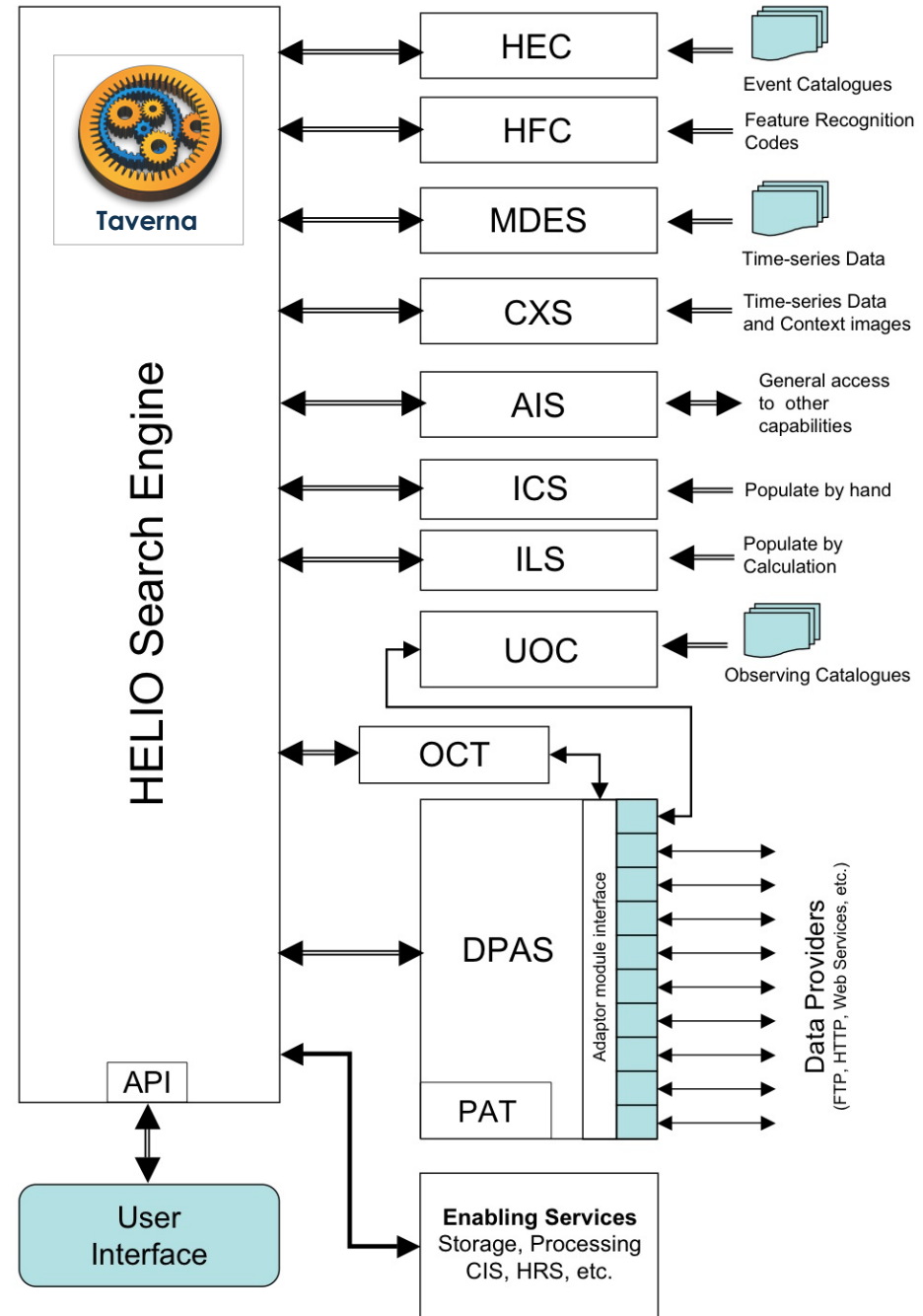
- Heliophysics Event Catalogue (HEC)
- Heliophysics Feature Catalogue (HFC)
- Metadata Evaluation Service (MDES)
- Context Service (CXs)
- Auxiliary Information Service (AIS)
- Propagation, etc. models*

- Match to observations

- Instrument Capabilities Service (ICS)
- Instrument Location Service (ILS)
- Observation Coverage Service (OCS)
  - Observation Coverage Table
  - Unified Observing Catalogue

- Locate and retrieve data

- Data Provider Access Service (DPAS)
  - Provider Access Table





## Defined Services

In addition to the services used to identify interesting events and then find and retrieve the data, there are a number of **Enabling Services**

These 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/>

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.

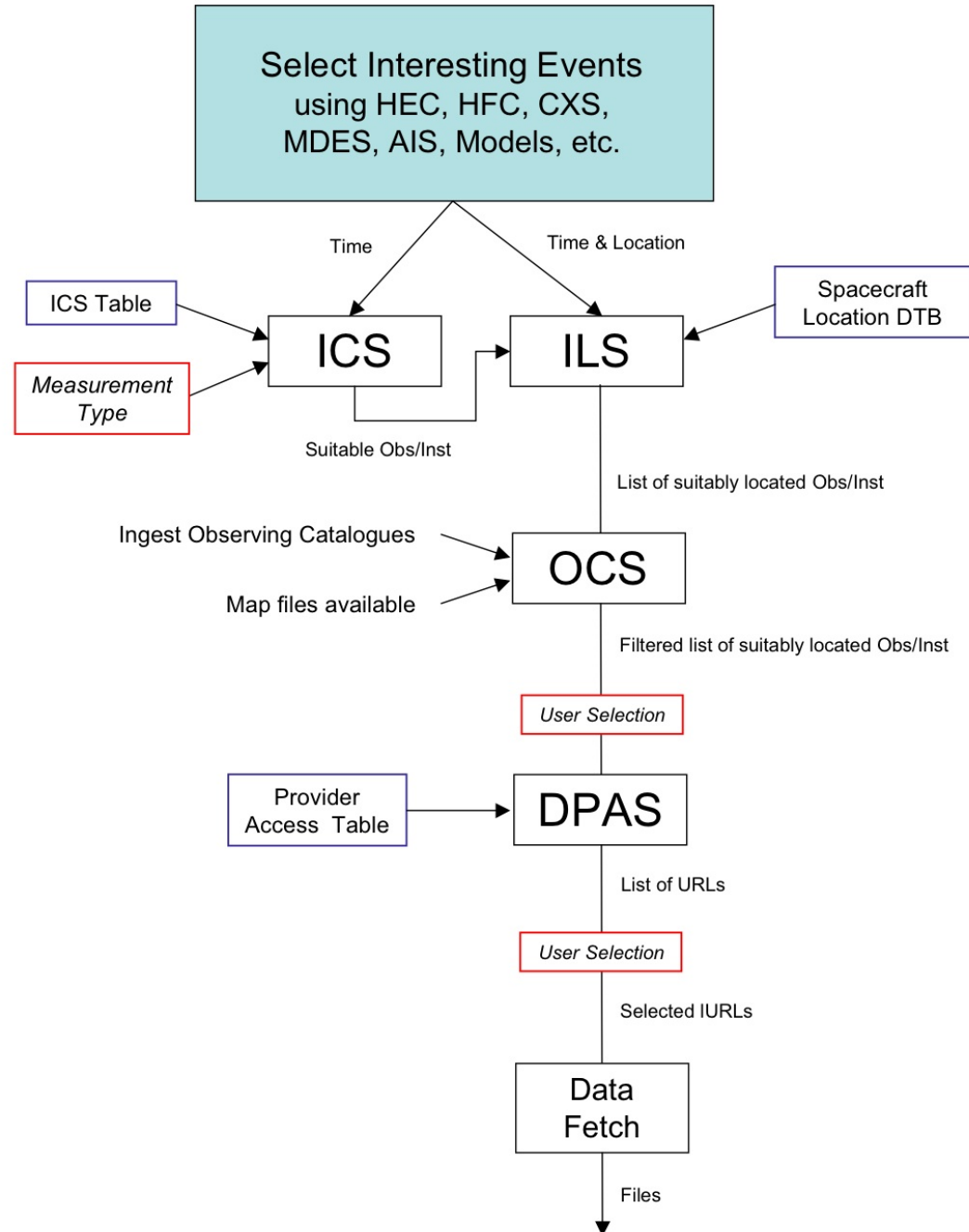




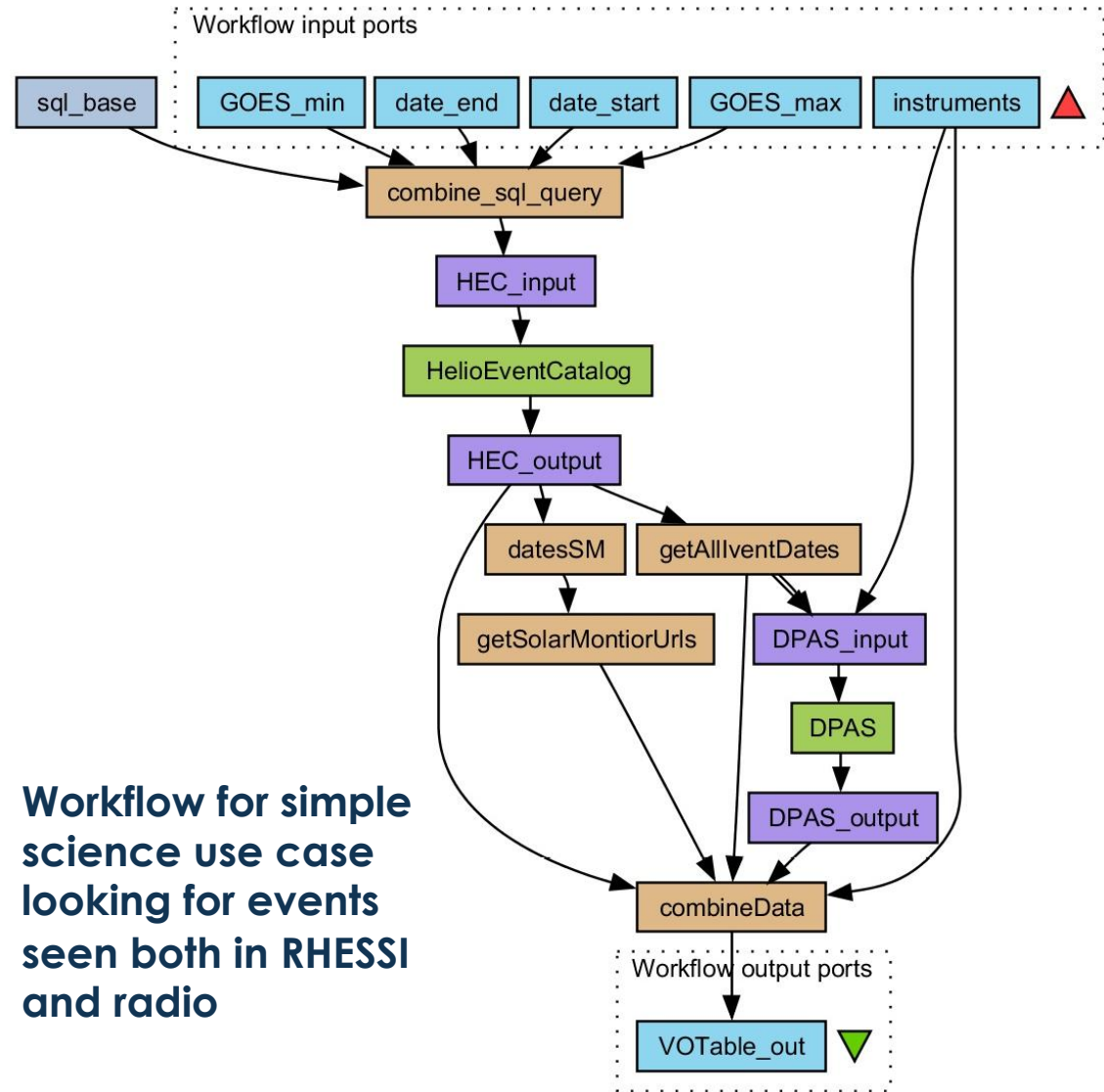


## Workflows...

- HELIO built from Services
- Services used individually or as part of a workflow
  - **Taverna** workflow tool base-lined for HELIO
  - Other workflow tools and scripting languages possible
- **Within event selection, the user may iterate using several services and tools until they are satisfied**
- **Other services then help the user go from a list of interesting times and locations to a list of Instruments and times to a list of files**



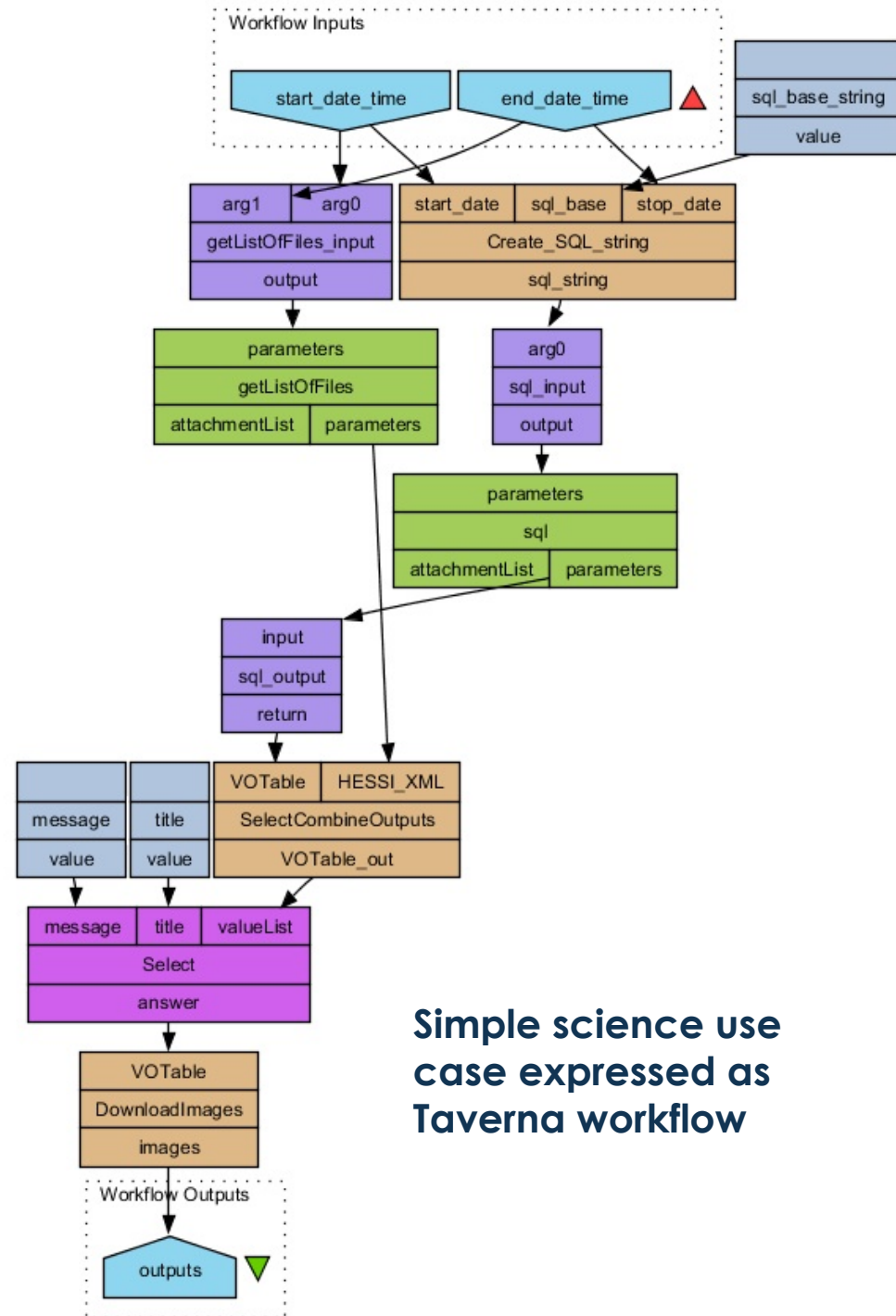
- **HELIO built from Services**
- **Services used individually or as part of a workflow**
  - **Taverna** workflow tool base-lined for HELIO
  - Other workflow tools and scripting languages possible
- **Within event selection, the user may iterate using several services and tools until they are satisfied**
- **Other services then help the user go from a list of interesting times and locations to a list of Instruments and times to a list of files**





# Workflows...

- HELIO built from Services
- Services used individually or as part of a workflow
  - Taverna workflow tool base-lined for HELIO
  - Other workflow tools and scripting languages possible
- Within event selection, the user may iterate using several services and tools until they are satisfied
- Other services then help the user go from a list of interesting times and locations to a list of Instruments and times to a list of files



Simple science use case expressed as Taverna workflow





# Status and Networking



- **Status**

- HELIO being developed as a series of Releases – starting R4
- Working prototypes of several service exist, some can be accessed
  - HEC, ICS & ILS, MDES, UOC, CXS, DPAS and HRS
  - Taverna Server now available ⇒ workflows “invisible to the user”
    - Hand-crafted Workflows executed on Workbench on user’s system until now
  - Simple user interface
- Lists of metadata and data that could be added has been defined
  - DPAS can already access data from >150 instruments

- **Networking**

- Trying to implement a series of **User Groups**
  - Users that will help define Use Cases and refine system requirements
  - Users that will test and validate the system (over the duration of the project)
- First CDAW planned for November 2010
  - Driving/prioritizing the selection of metadata, data and services





# Outline



- **Overview of the types of data providers**
  - Review of some European VOs
- **Description of the Heliophysics Integrated Observatory**
- **Standards and interoperability**
  - Issues related to use of data by HELIO
  - Description of CASSIS
- **Use Cases**
- **Summary**





# Standards and Interoperability



- **To achieve its objectives, HELIO must address a number of issues**
  - These are what make everything possible
- **Must support a search across domains 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
  - Need to annotate the metadata to properly describe quantities
    - Unambiguously define the parameters involved
    - Put them in context of other information that are used
- **Must provide integrated access to data from many domains**
  - Solar, heliospheric, planetary, geophysics (magnetospheric, ionospheric)
    - Different file formats, ways of storing, handling and using the data
  - Variety of access techniques
    - Access protocols – http, ftp, cg-bin, Java-script, Web-service, ...





# Standards and Interoperability



- **Developing data models to help describe the systems that we are trying integrate**
  - Based on models from EGSO, SPASE, etc.
  - Trying to create a data model that over-spans the others...
- **Developing standards for files that we are producing that could be extended to other groups**
  - Annotation also important in these
- **Developing recommendation for providers about how to naming and storage strategies that would make their data archives more interoperable**
  - Activities related to IAU working group and CASSIS
- **Potential users of all VOs, research infrastructures and data providers should be aware of the issues and lobby to ensure that things are improved!!**



List of Interplanetary Coronal Mass Ejections (ICMEs) Observed by STEREO A/B										
#	STEREO	Start time [Year Doy Month/Day HH:MM]	Magnetic obstacle (~ flux rope) start time	End time	Ptmax [pPa]	Bmax [nT]	Vmax [km/s]	$\Delta V^1$ [km/s]	Group <sup>2</sup>	Comment
<b>2006</b>										
1	A & B	2006 348 12/14 14:12	2006 348 12/14 23:04	2006 349 12/15 14:08		18 (20)				plasma data data gap, like Group 3 event

**The STEREO event lists for IMPACT have problems:**

- Very different ways of expressing time
- Location of spacecraft not explicitly included
- No information on “local” solar B angle...

**Good for creativity; very poor for interoperability**

#	STEREO	Start time [Year Doy Month/Day HH:MM]	Magnetic obstacle (~ flux rope) start time	End time	Ptmax [pPa]	Bmax [nT]	Vmax [km/s]	$\Delta V^1$ [km/s]	Group <sup>2</sup>	Comment	Hz data availability <sup>5</sup>	
											Y	F
1	A	2007 142 5/22 14:12									Y	F
2	B	2007 142 5/22 03:10									N	F
3											N	F
											N	forming F, gradual B increase
											3.91	1.4
											1.67	1.22
											4.35	1.53
<b>2007</b>												
1	1	B	63 03/04 18:50	66 03/07 20:00	65 03/06 06:50	150	10.1	21.4	343	635		
	1	A	63 03/04 15:50	66 03/07 20:00	65 03/06 10:40	85	10.8	16.9	352	632		
2	2	B	70 03/11 08:40	73 03/14 06:10	71 03/12 13:00	260	15	57.5	287	734		
	2	A	70 03/11 03:50	73 03/14 06:30	70 03/11 16:30	115	11.5	36.8	322	722		
3	3	B	82 03/23 08:30	87 03/28 12:00	82 03/23 19:40	170	14.3	41.3	277	621		
	3	A	82 03/23 03:10	87 03/28 12:00	83 03/24 20:10	100	12.1	40.8	280	578		





# Annotation of data that are used/created



```

<DESCRIPTION>Helio Sec time based query</DESCRIPTION>
<TABLE nrows="50">
<FIELD arraysize="*" datatype="char" name="goes_id">
<DESCRIPTION>Sec Goes Id</DESCRIPTION>
</FIELD>
<FIELD arraysize="*" datatype="char" name="ntime_start">
<DESCRIPTION>N Time Start</DESCRIPTION>
</FIELD>
<FIELD arraysize="*" datatype="char" name="time_start" ucd="time.start">
<DESCRIPTION>Start Time</DESCRIPTION>
</FIELD>
<FIELD arraysize="*" datatype="char" name="time_peak" ucd="time.phase">
<DESCRIPTION>Peak Time</DESCRIPTION>
</FIELD>
<FIELD arraysize="*" datatype="char" name="time_end" ucd="time.end">
<DESCRIPTION>End Time</DESCRIPTION>
</FIELD>
<FIELD arraysize="*" datatype="char" name="nar">
<DESCRIPTION>Active Region</DESCRIPTION>
</FIELD>
<FIELD arraysize="*" datatype="char" name="ntime_end">
<DESCRIPTION>N End Time</DESCRIPTION>
</FIELD>
<FIELD arraysize="*" datatype="char" name="latitude" ucd="pos.solar.lat">
<DESCRIPTION>Latitude</DESCRIPTION>
</FIELD>
<FIELD arraysize="*" datatype="char" name="longitude" ucd="pos.solar.long">
<DESCRIPTION>Longitude</DESCRIPTION>
</FIELD>

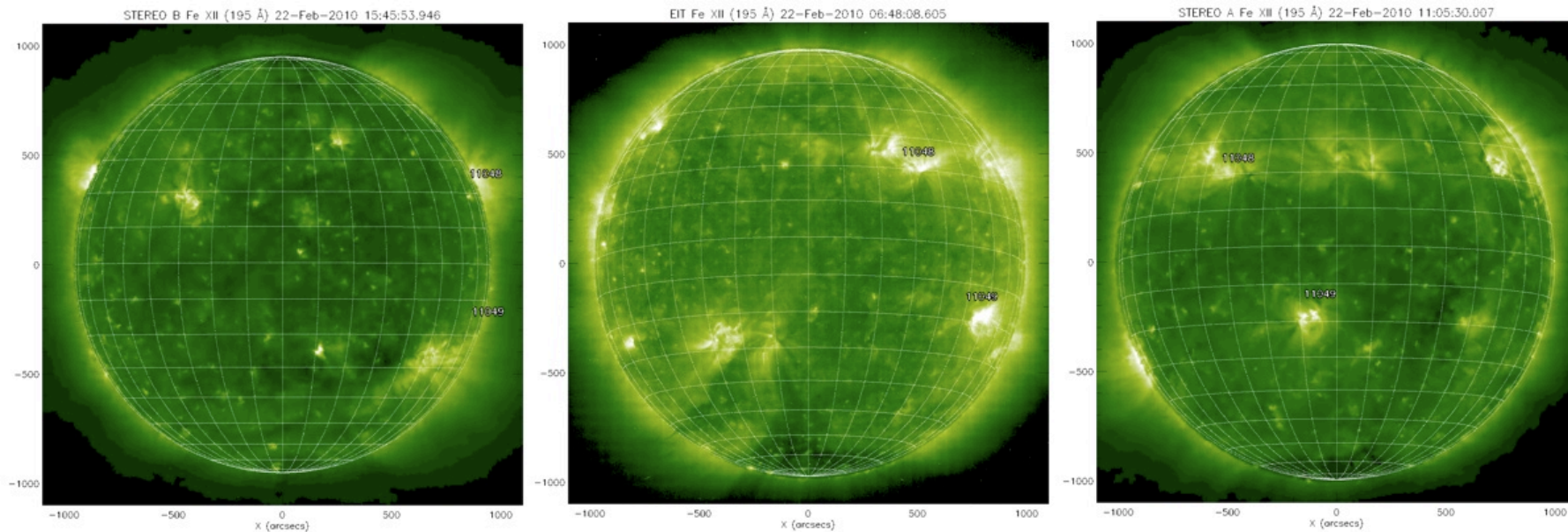
```

S   pos.cartesian	Cartesian (rectangular) coordinates
Q   pos.cartesian.x	Cartesian coordinate along the x-axis
Q   pos.cartesian.y	Cartesian coordinate along the y-axis
Q   pos.cartesian.z	Cartesian coordinate along the z-axis
S   pos.cmb	Cosmic Microwave Background reference frame
Q   pos.dirCos	Direction cosine
V   pos.distance	Linear distance
S   pos.earth	Coordinates related to Earth
Q   pos.earth.altitude	Altitude, height on Earth above sea level
Q   pos.earth.lat	Latitude on Earth
Q   pos.earth.lon	Longitude on Earth
S   pos.ecliptic	Ecliptic coordinates
Q   pos.ecliptic.lat	Ecliptic latitude
Q   pos.ecliptic.lon	Ecliptic longitude
S   pos.eop	Earth orientation parameters
Q   pos.eop.nutation	Earth nutation
Q   pos.ephem	Ephemeris
S   pos.eq	Equatorial coordinates
Q   pos.eq.dec	Declination in equatorial coordinates
Q   pos.eq.ha	Hour-angle
Q   pos.eq.ra	Right ascension in equatorial coordinates

- **UCDs and utypes** are concepts defined by the **International Virtual Observatory Alliance (IVOA)**
- Existing UCDs, etc not really applicable to heliophysics; HELIO needs to develop a set!

- **Outputs from the HEC and other services are in VOTable format**
- **HELIO has decided that lists that could be used by external communities should be fully annotated with UCDs and utypes – also useful internally!**
  - **Unambiguously defines what the quantities are and mean**
- **Need to develop a comprehensive, over-spanning data model**
  - **Looking for help to define utypes and UCDs for the Solar System**





STEREO-B, SOHO and STEREO-A on 22/2/2010 when the STEREO spacecraft were  $\sim 65^\circ$  behind/ahead of Earth

- **HELIO must use data from many points in the Solar System**
- **The same feature can be seen from several different vantage points**
  - Need to be able to relate features seen in pairs of images
  - Need to be able to interpret coronagraph (plane-of-sky) observations from different viewpoints and relate them to motion of material in the heliosphere
  - Need to relate these to *in-situ* observations made at various locations
  - Need to be able to track features and follow effects into planetary environments
- **Working set of coordinates has been selected for developing the system...**



## Some simple guidelines for Providers...



- **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
- **File Metadata** – It is essential that all files contain good metadata describing in detail how the observations were made
  - May be impossible to use the data in some circumstance if the metadata are not properly formed (files need to be FITS or equivalent)
- **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



- **CASSIS is a Coordination Action 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 domains**
  - CASSIS brings together three projects that are directly relevant to this issue
  - 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 disciplines**
  - Many aspects of the system are inter-related
  - Difficult to address them because of the lack of the integrating tools and techniques
- **Advances in technology means that the intrinsic differences between disciplines are being addressed**
  - Manifest by differing data formats and other dependencies
- **Three FP7 projects – HELIO, EuroPlanet RI & SOTERIA.**
  - 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 intended to take things to the next level by cooperating in a number of areas**
  - Enabling new combinations of interdisciplinary studies

- **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



# Outline



- **Overview of the types of data providers**
  - Review of some European VOs
- **Description of the Heliophysics Integrated Observatory**
- **Standards and interoperability**
  - Issues related to use of data by HELIO
  - Description of CASSIS
- **Use Cases**
- **Summary**





# Use Cases and Capabilities



- **We will examine various use cases and discuss how the search capabilities of HELIO can be used to address them**
  - **Basic Use Cases**
  - **Use Cases – Increased complexity**
  - **Use Cases – Propagating Phenomena**
  
  - **Instrument Capability and Location Services (ICS & ILS)**
  - **Heliophysics Event Catalogue (HEC)**
  - **Metadata Evaluation Service (MDES)**
  - **Ability to run and incorporate propagation models**
- **The capabilities provide by HELIO are an extension of those that are required to study Space Weather (SWx++)**







## Basic Use Cases



- **Simplest data request**
  - Find data where user knows instruments and times (and locations)
- **More complex data request**
  - Find observations of a particular types where user knows times and locations but not the instruments
  - Need to determine which instruments with the required capabilities are located in the right place a the right time





## 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.





## 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:**
  - **Calculate the location of satellites in the orbits around a planet**
    - **Location with respect to magnetospheric or ionospheric boundaries needs more information**
  - **Determine whether the observatory lies within a region-of-interest**
    - **Location of the propagating phenomena determined by modelling**





## Basic Use Cases – Using selected parts of HELIO



- **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
  - [Observation Coverage Table (OCT) used to help Provider Access Service (DPAS) determine availability of 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**
  - ICS can be used to identify suitable instruments
    - Also approximate location if they are on/near a planet
  - 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-driven 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
    - Seen at earth; seen elsewhere
  - Disturbance in the solar wind
    - Seen at Earth – (geo) magnetic storm
    - 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
    - Event and Feature Catalogues (HEC and HFC), Metadata Evaluation Service (MDES) [and Context Service (CXS)]
    - **Search may not follow a pre-defined workflow**
  - Results can be used as input in search for useful observations
    - Information can be stored on the system if user has authenticated





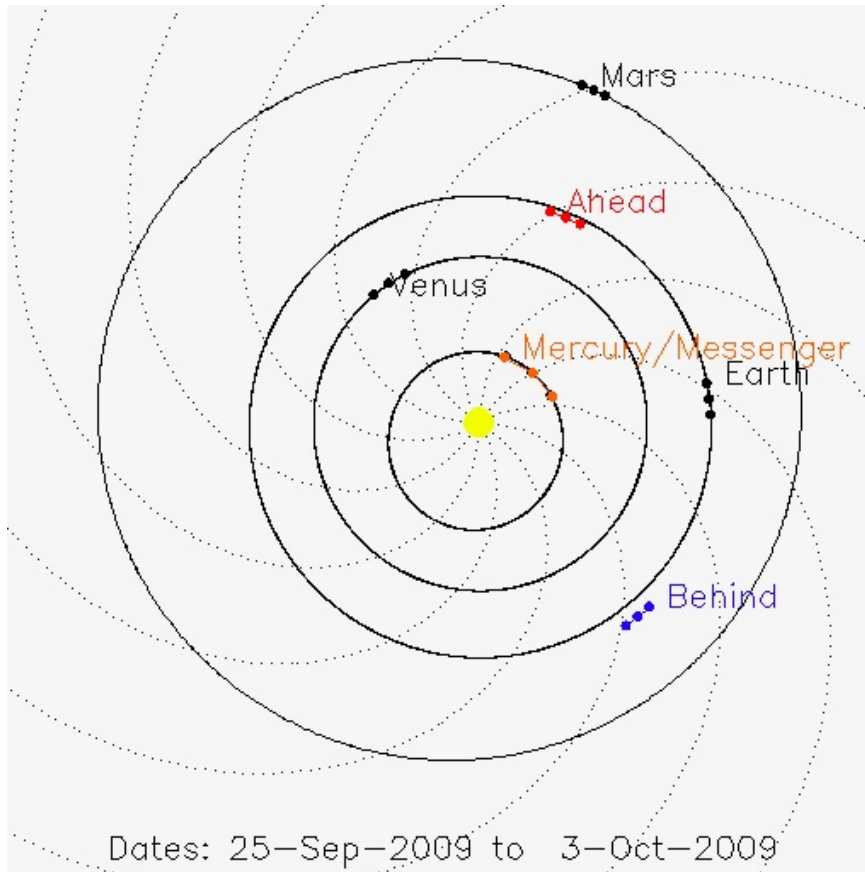
# HEC and MDES



- **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 user 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)

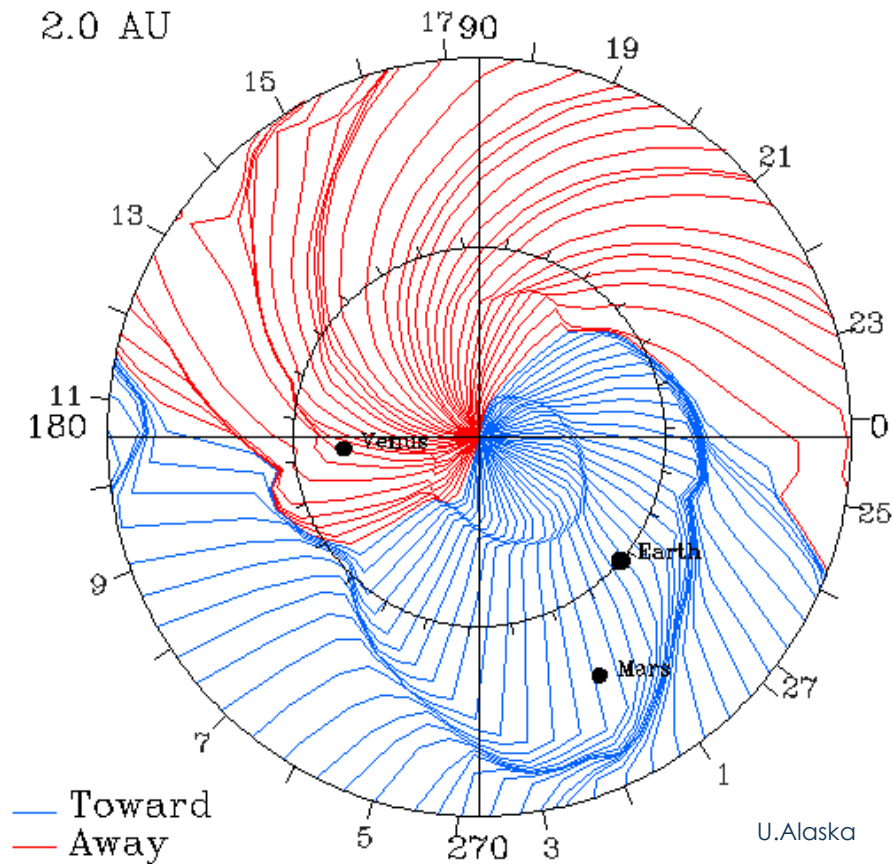


Table name	List Name	Description/URL
<i>Primary Solar Event Lists</i>		
goes_xray_flare	GOES EVent list (GEV)	Edited list of GOES flares, locations from SXT, EIT, SXI (NGDC) <a href="#">[URL]</a>
halpha_flares_event	Solar H-alpha Flare events	List of H-alpha flares assembled from ground-based observatories (NGDC) <a href="#">[URL]</a>
sgas_event	SGAS Energetic Events	Energetic Solar Events, plus radio observations (NOAA) <a href="#">[URL]</a>
lasco_cme_cat	LASCO CME Catalogue	Edited list of CMEs detected by SOHO/LASCO (NASA/GSFC) <a href="#">[URL]</a>
hi_cme_list	STEREO HI CME list	List of CMEs detected by STEREO/HI (SSTD) <a href="#">[URL]</a>
<i>Geo-effective Events – Related to energetic particles</i>		
noaa_proton_event	Solar Proton Events	Geo-effective Solar Energetic Proton (SEP) events (NOAA) <a href="#">[URL]</a>
gle_list	Ground Level Enhancements	Enhancements detected by ground-based neutron monitors (Bartol) <a href="#">[URL]</a>
forbush_list	Forbush Decrease Events	Forbush Decreases, related to CMEs (Contrib.: E. Eroshenko, Izmiran) <a href="#">[URL]</a>
<i>Geo-effective Events – Related to variations in the Solar Wind</i>		
ssc_list	Storm Sudden Commencements	Indicator of the onset of a Geo-magnetic storm (NGDC) <a href="#">[URL]</a>
aastar_list	Major Magnetic Storms (AA*)	Major Magnetic Storms according to the AA* criteria (NGDC) <a href="#">[URL]</a>
apstar_list	Major Magnetic Storms (Ap*)	Major Magnetic Storms according to the Ap* criteria (NGDC) <a href="#">[URL]</a>
bas_magnetic_storms	BAS Magnetic Storms	Geo-effective magnetic storms, Dst < -50 nT (Contrib.: R. Horne, BAS) <a href="#">[URL]</a>
<i>General events in the Solar Wind</i>		
solar_wind_event	Solar Wind Events	Events in the solar wind observed by Wind, ACE and Imp 8 (NASA) <a href="#">[URL]</a>
<i>Solar Event Lists from individual observatories (flares, etc.)</i>		
yohkoh_flare_list	Yohkoh HXT Flare List	Flare list from the Yohkoh Hard X-ray Telescope <a href="#">[URL]</a>
yohkoh_sxt_trace_list	Yohkoh SXT TRACE flare list	Common flares observed by Yohkoh/SXT and TRACE <a href="#">[URL]</a>
hessi_flare	RHESSI Flare List	List of hard X-ray flares seen by RHESSI <a href="#">[URL]</a>
kso_flare	Kanzelhoehe Flare List	Flares observed in H-alpha by the Kanzelhoehe Observatory <a href="#">[URL]</a>
eit_list	EIT Waves	Waves seen in the EUV by SOHO/EIT (Contrib.: B. Thompson, NASA/GSFC) <a href="#">[URL]</a>
lasco_cme_list	LASCO Preliminary CME List	List of CMEs detected by SOHO/LASCO built from Alerts (SOHO/GSFC) <a href="#">[URL]</a>
<i>Summary Data</i>		
srs_list	Solar Region Summary	Parameters describing active regions: lat, long, size... (NOAA/USAF) <a href="#">[URL]</a>
dsd_list	Daily Solar Data	Parameters describing solar activity: SSN, flare types... (NOAA) <a href="#">[URL]</a>

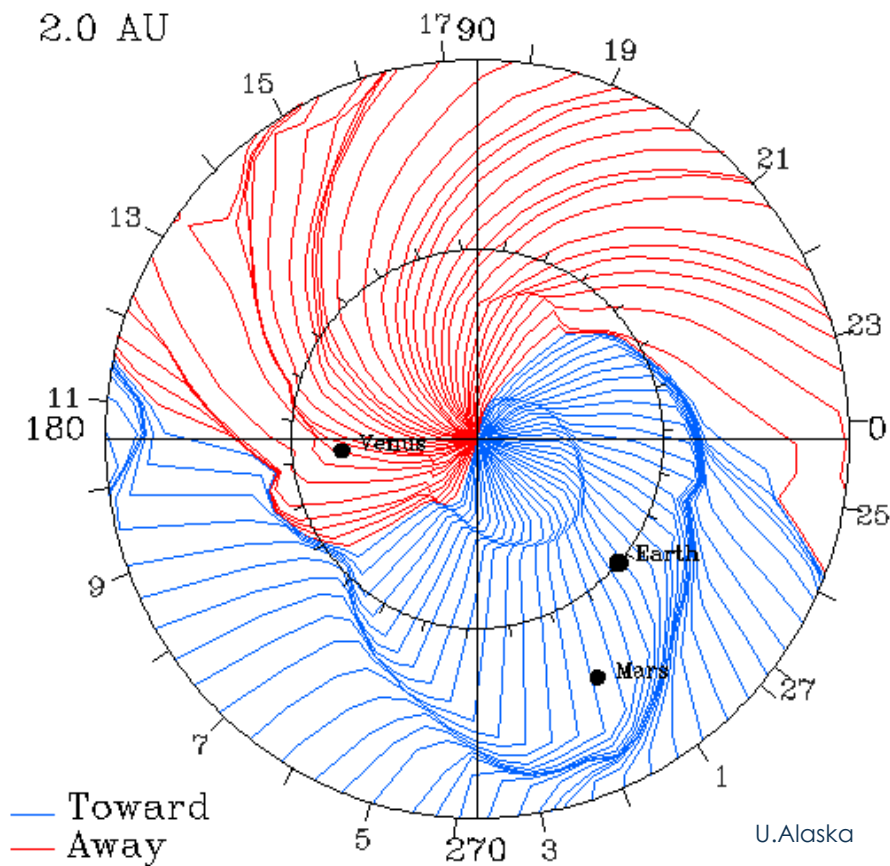


- User has identified an event in one part of the solar system and is needs related observations
  - Case: Messenger fly-by of Mercury on 29 September 2009
- Simple model of Parker Spiral indicates which observatories are on same or close field lines
  - User employs model to determine delays related to effects
- ICS/ILS and OCS determine which instruments probably made observations
- DPAS used to locate and retrieve the required observations



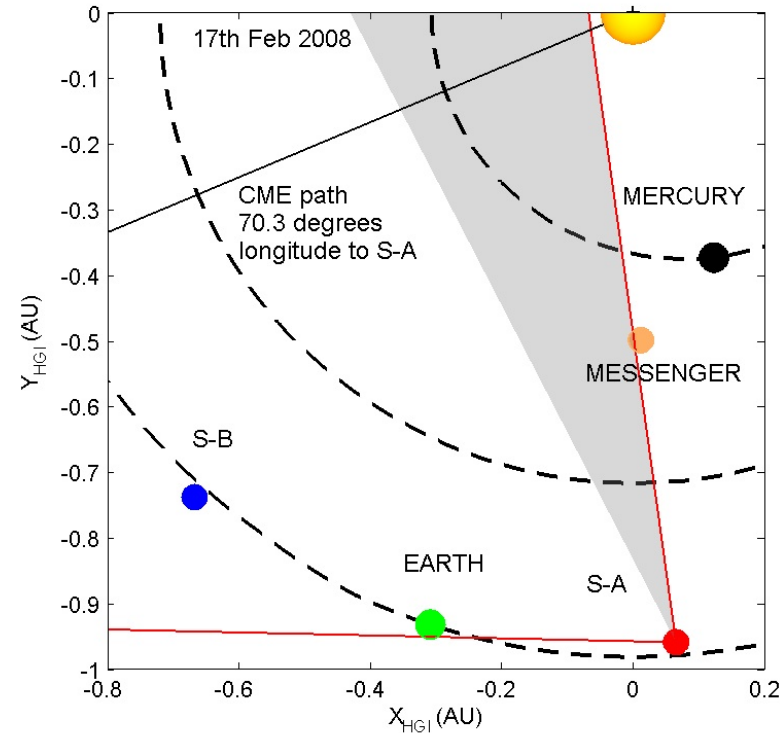
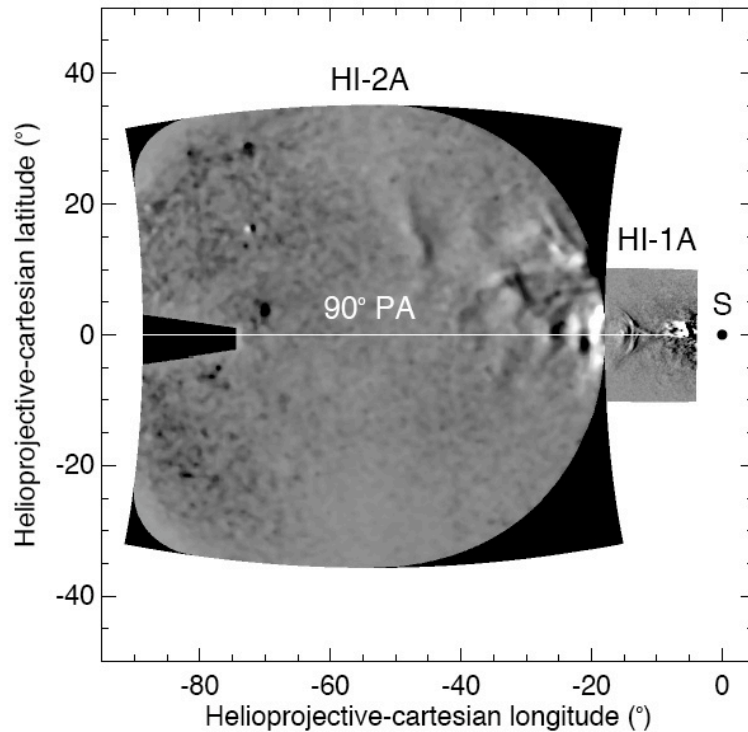


- User wishes to examine effects of a CME as it propagates through the Solar System
- HEC used to identify time and location of events on the Sun
- Propagation model used to determine times & locations in other parts of the Solar System
- ICS/ILS identify instruments that were suitably located to make 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



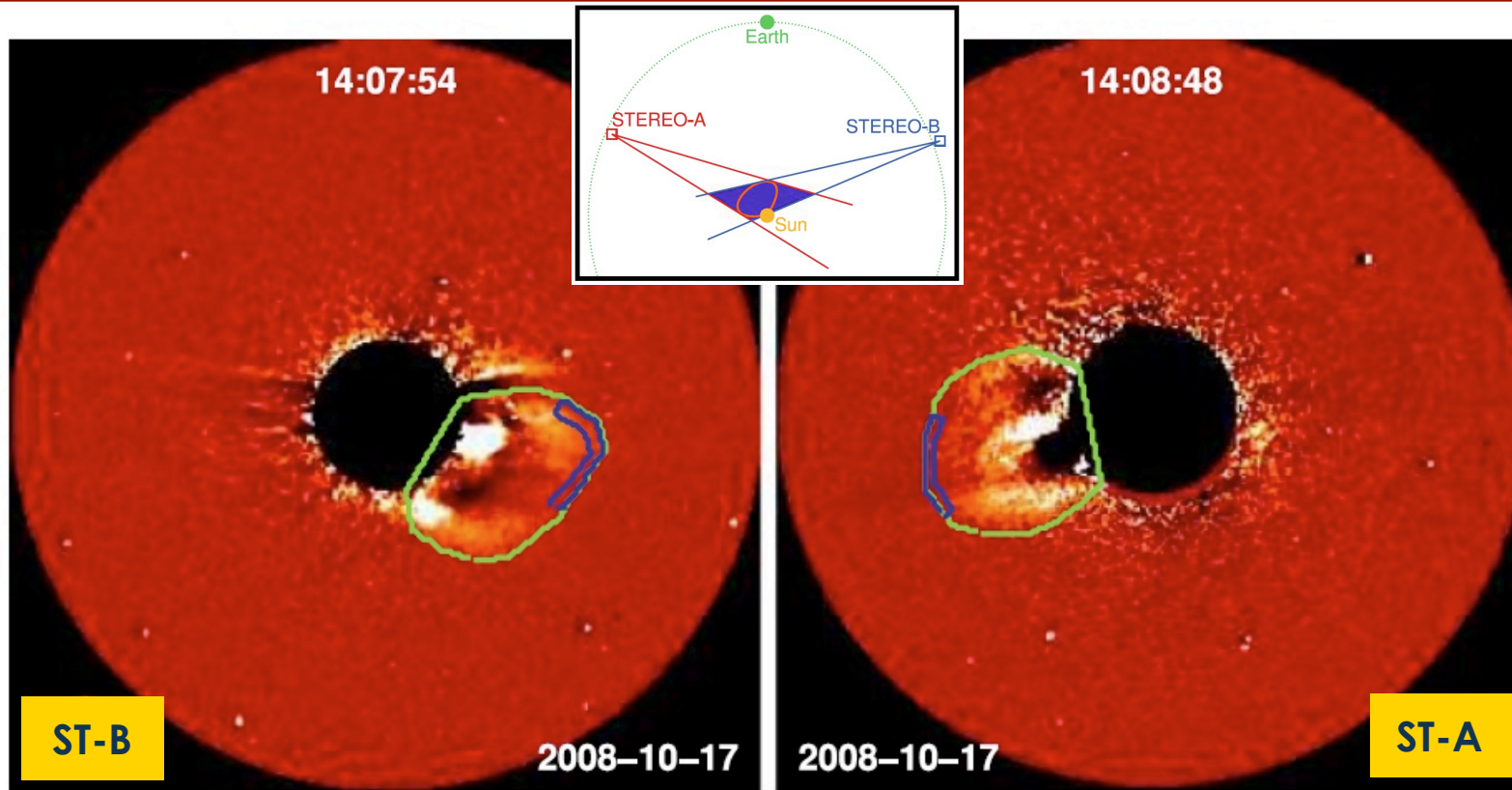
- User knows of a space weather effect at the Earth and wants to examine the possible
- Propagation model used to determine times & locations back at the Sun
- Search tools aid search
  - HEC used to identify events
  - HFC used to identify relevant features – filaments, CHs
  - MDES used to look for 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

08:09 UT, 17 February 2008



Savani et al, 2009

- **HELIO's search capability allow the user to find various types of observations that are relevant to their science goals**
- **User then needs to understand what they are seeing**
  - *Interpretation of plane-of-sky information is an issue...*



Konig et al., 2009

- Once the data have been interpreted, they can be used to refine the propagation model
- The process can be iterative...



# Use Cases drive system development



- **By examining use cases we derive criteria for**
  - Developing additional system capabilities
  - Adding different types of metadata, data and derived products
- **As the system grows we may need to enhance capabilities that have already been deployed**
  - Instrument Location Service (ILS) needs further development to support more complex queries
  - Selection of data that DPAS returns pointers to needs further thought
    - Shear quantities of data that can be accessed is a real problem!!
- **Always on the look-out for types of information that will allow further growth**
  - Shopping list for required sets of information – e.g. for the HEC





## HEC – Desired solar event lists



- Whether event is seen depends on location of observer w.r.t. source
- Flare events (remote)
  - Flare observed from near Sun-Earth line
    - GOES x, compound H-alpha, plus individual observatory lists
  - Flares observed from other vantage points
    - STEREO-EUVI, ?
- Other types of surface/atmospheric events and features (remote)
  - Waves, filaments, coronal holes, etc...
- CME-onset events (remote)
  - CMEs observed from near Sun-Earth line
    - SOHO-LASCO plus ground-based coronagraphs
  - CMEs observed from other vantage points
    - STEREO-COR2 CME event, -HI events
  - Interpretation of plane-of-sky information is an issue...
- Perturbation in the Solar Wind (remote)
  - Emissions related to propagating phenomena
    - Heliospheric imagers (SOHO-SWAN, SMEI, STEREO-HI)
    - Radio observations...





## HEC – Desired *in-situ* event lists



- ***In-situ* observations are of the environment around the observer**
  - Observer experiences passing phenomena...
- **Effects could be related to CMEs, SIRs, sector boundaries, etc.**
  - Whether observed depends on observer's location wrt point of origin
  - Effects modified outside/inside of planetary magnetosphere
- **Particle lists (*in-situ*)**
  - Events observed near Earth
    - GOES x, ACE..., plus ground-level events
  - Events observed elsewhere
    - STEREO, Ulysses, planetary orbiters...
- **Perturbation in the Solar Wind (*in-situ*)**
  - Events observed near Earth
    - SOHO, ACE, WIND, IMP8...
  - Events observed in the heliosphere and environs of other planets
    - STEREO, Ulysses, planetary orbiters
- **Galactic cosmic-ray (GCR) flux modified by structures in heliosphere**





# Summary



- **Integrated access to all types of observational data becoming increasingly important to support scientific analysis**
- **We have to find ways into the increasing volumes of information that are focused on scientific need and that isolate the user from foibles of the systems that provide access to the data**
- **Virtual Observatories play a key roll because they bring considerable added value to the process**
  - **HELIO making progress; there is a still a lot to do**
- **Improved interoperability would help us all**
  - **Quality of data and metadata is a concern to to us all**
  - **Providers have common problems even if they do not realize it**
    - **Providers of all types are important to give complete picture**
  - **User involvement in discussion is important to the process**
    - **CASSIS intended to facilitate this – everyone is welcome...**







