



Services and Workflows for the Heliophysics Integrated Observatory

André Csillaghy, David Guevara, Marco Soldati
Institute of 4D Technologies, University of Applied Sciences
North Western Switzerland (FHNW)

Robert D. Bentley
Mullard Space Sciences Laboratory, University College London

Jean Abouadarham
LESIA, Observatoire de Paris Meudon

Cristian Jacquey
CDPP/CESR, Université Paul Sabatier Toulouse

Gabriele Pierantoni
Computer Science Department, Trinity College Dublin

Anja LeBlanc, John Brooke
Computer Science Department, University of Manchester

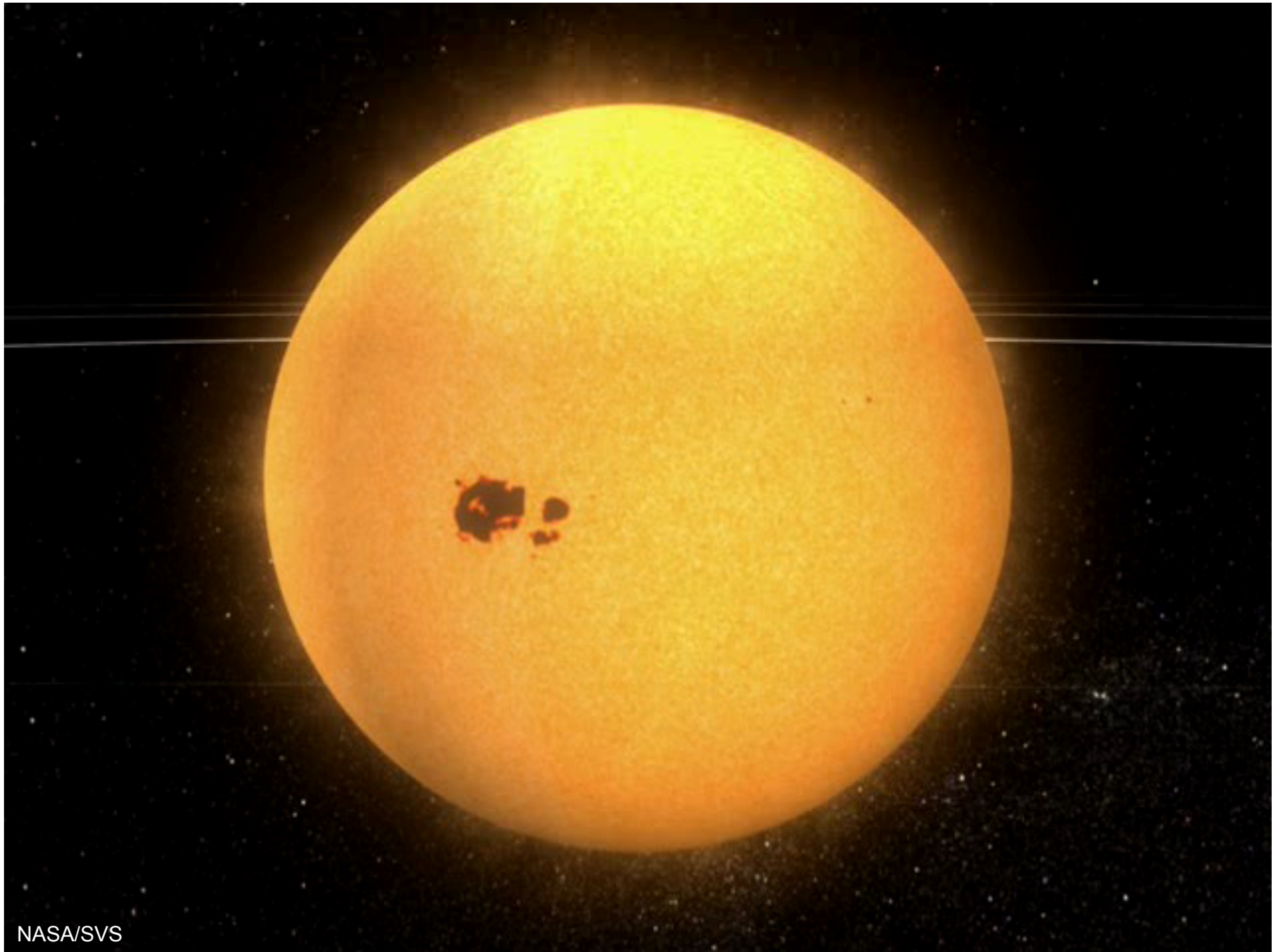
for the HELIO Team

www.helio-vo.eu



Heliophysics Integrated Observatory (HELIO)

CONTEXT



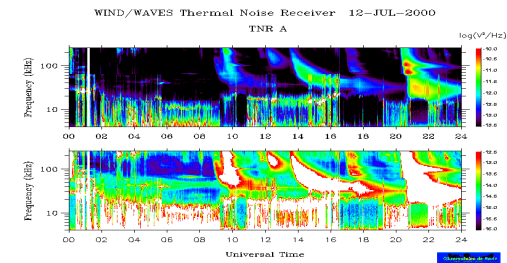
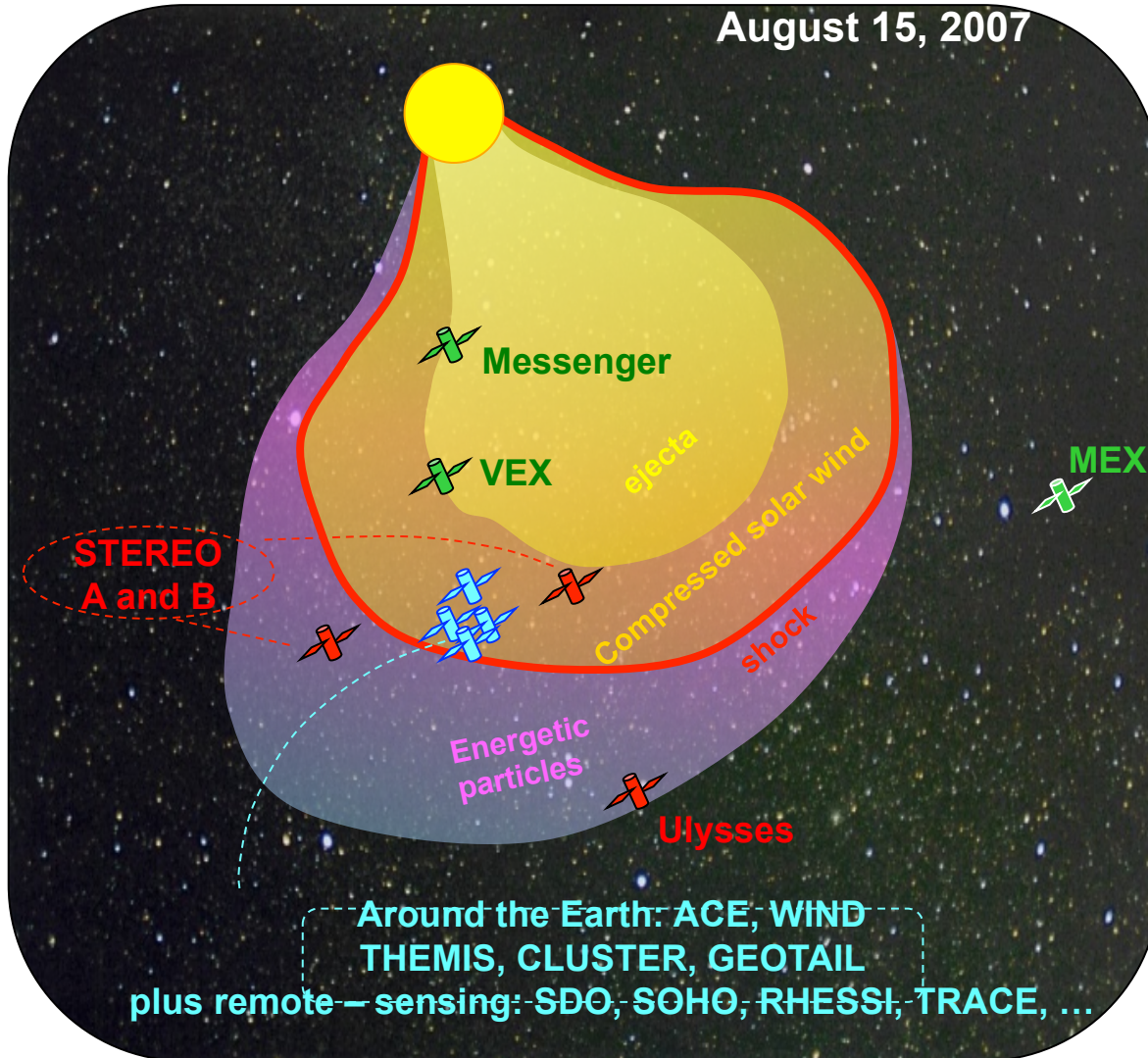
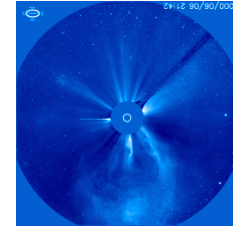
NASA/SVS

You need a bunch of spacecraft to study CMEs

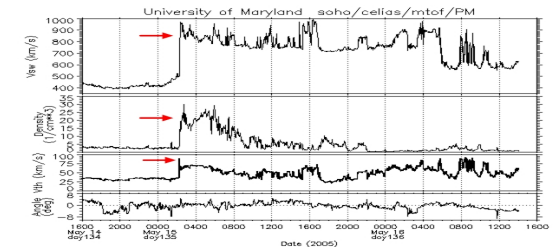
- %Origin and propagation?
- %Energetic particle acceleration?
- %Impact of the Earth environment?

Needs:

- Solar data
- Remote sensing from radio data

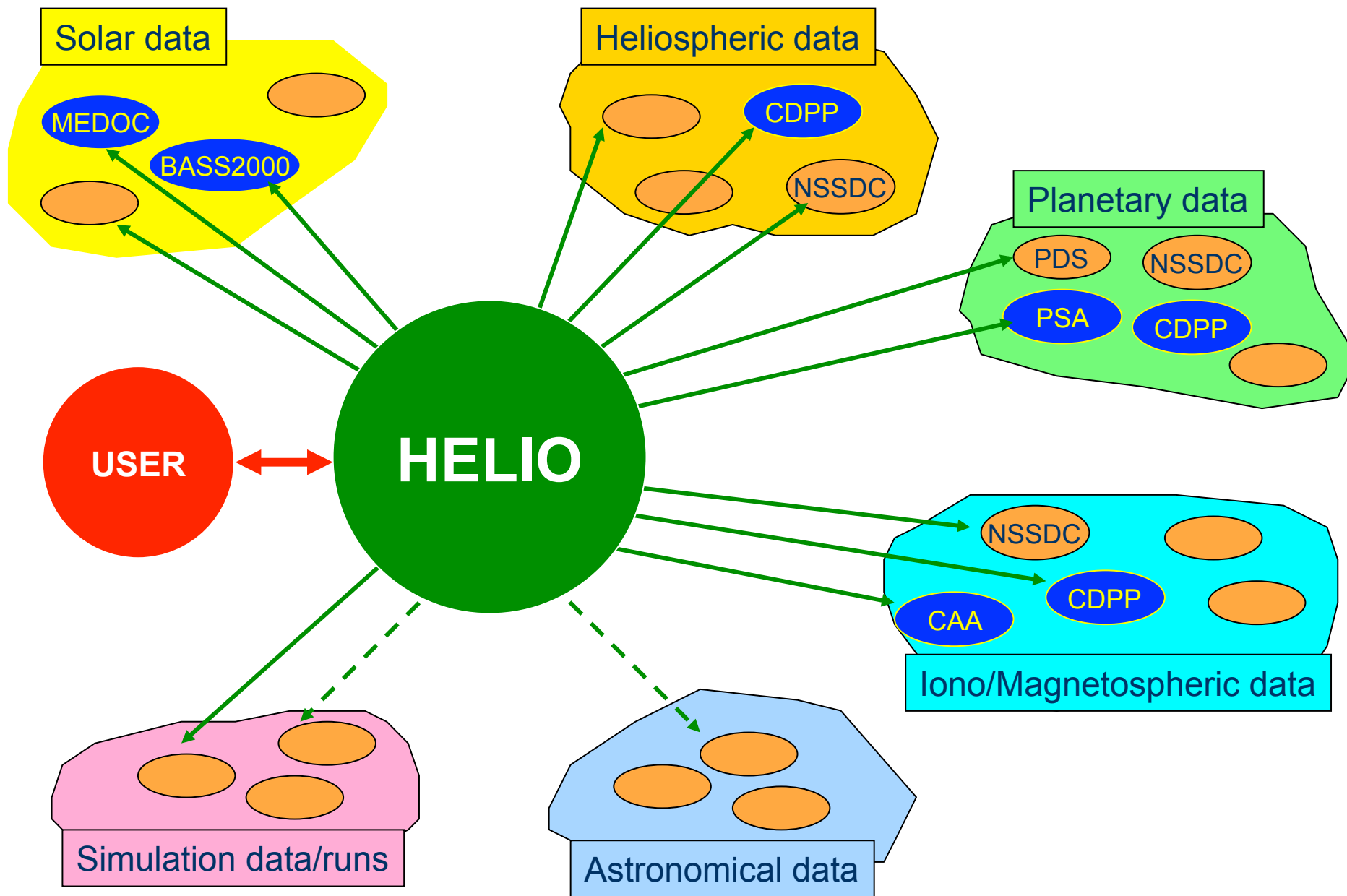


- In situ measurements



- Model runs or simulation results

And services for performing their integrated analysis





What is HELIO?



- **HELIO is a software infrastructure to**
 - Create a **collaborative environment** where scientists can study the connection between solar phenomena, interplanetary disturbances and their effects on the planets (especially the Earth)
 - Establishing **standards** for accessing heliophysical data following as much as possible already established rules (e.g. IVOA standards)
- **HELIO must match the needs of the community, however it chooses to use it.**
 - “Agile Development”: Small but frequent iterations
- **HELIO integrates solar and heliospheric (meta) data**
 - Essentially a metadata system (not a data center)
- **HELIO is funded under EC’s FP7 Capacities Specific Program “Scientific Data Infrastructures”**
 - Started 1 June 2009, duration 36 months

SERVICES

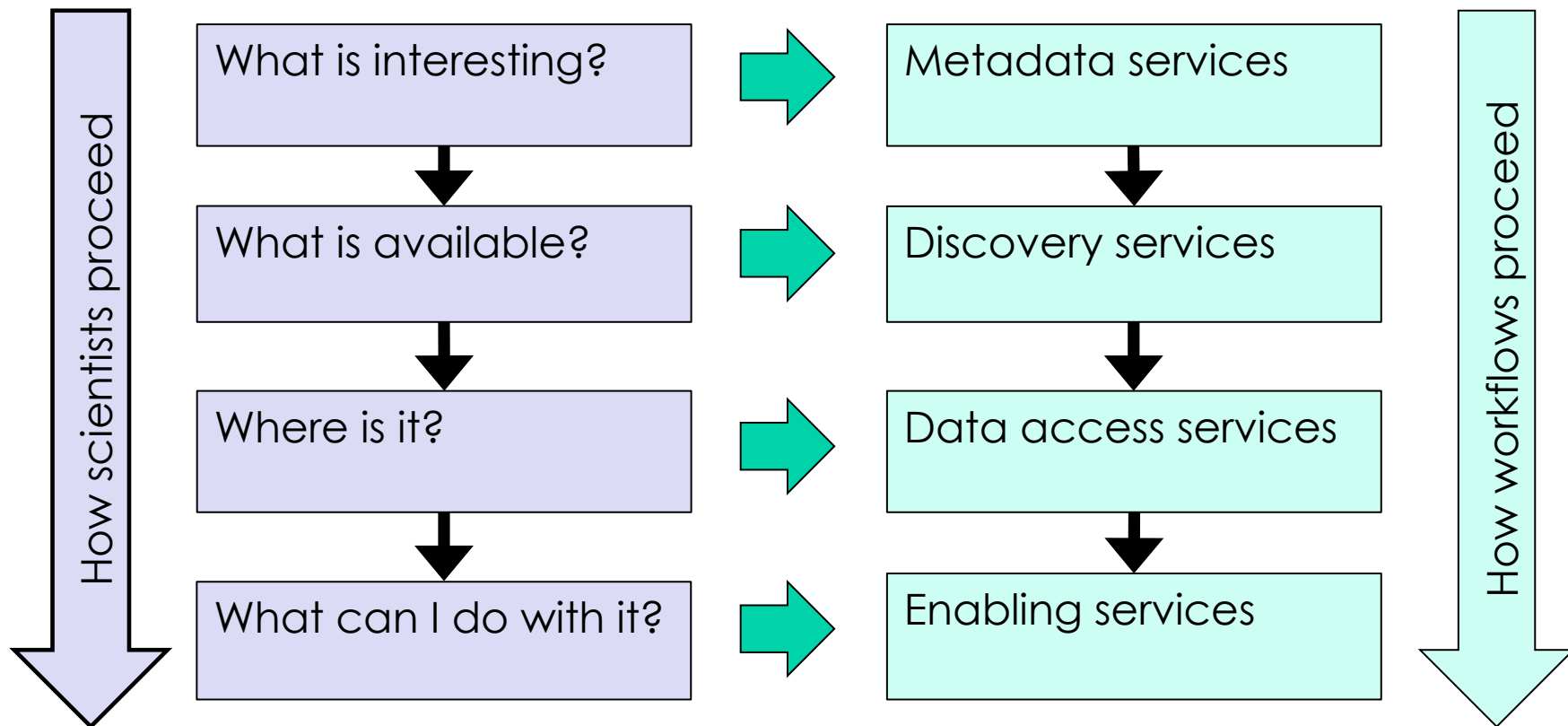


HELIO's perspective of a generic, time-driven problem



- **Identify interesting things to study**
 - *Search based solely on metadata and derived products*
 - Event lists and feature lists used as primary selection criteria
 - *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
- **Review availability of suitable observations**
 - Determine whether suitable instruments are at the relevant locations
 - Science objectives dictate types of observations required
 - Determine whether instrument was making observations
 - Coverage and quality of observations are additional 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
- **Do something with the data**

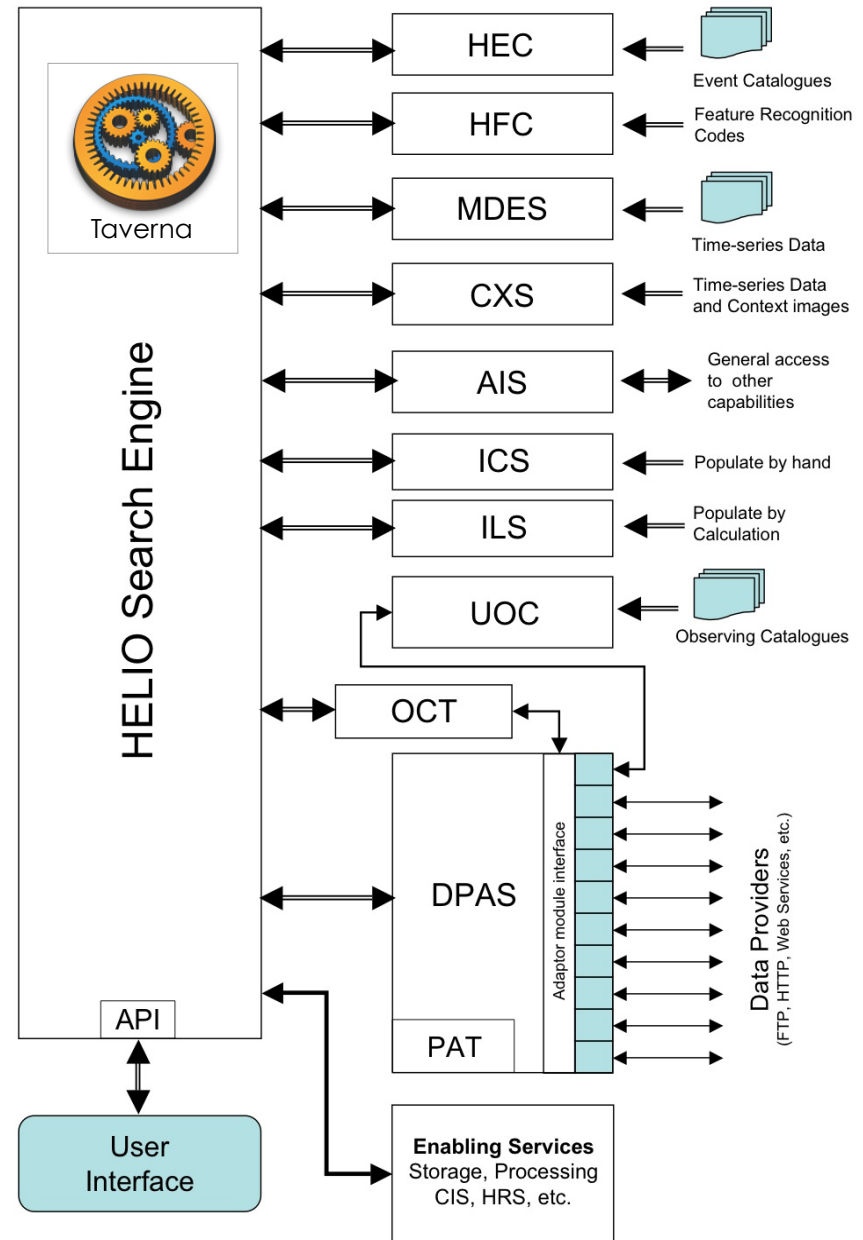
- **Exploit the generic aspect of the basic problem, to reuse software components over and over again**
- **“Translate” the generic problem into technical terms**





Putting things together: The HELIO SOA

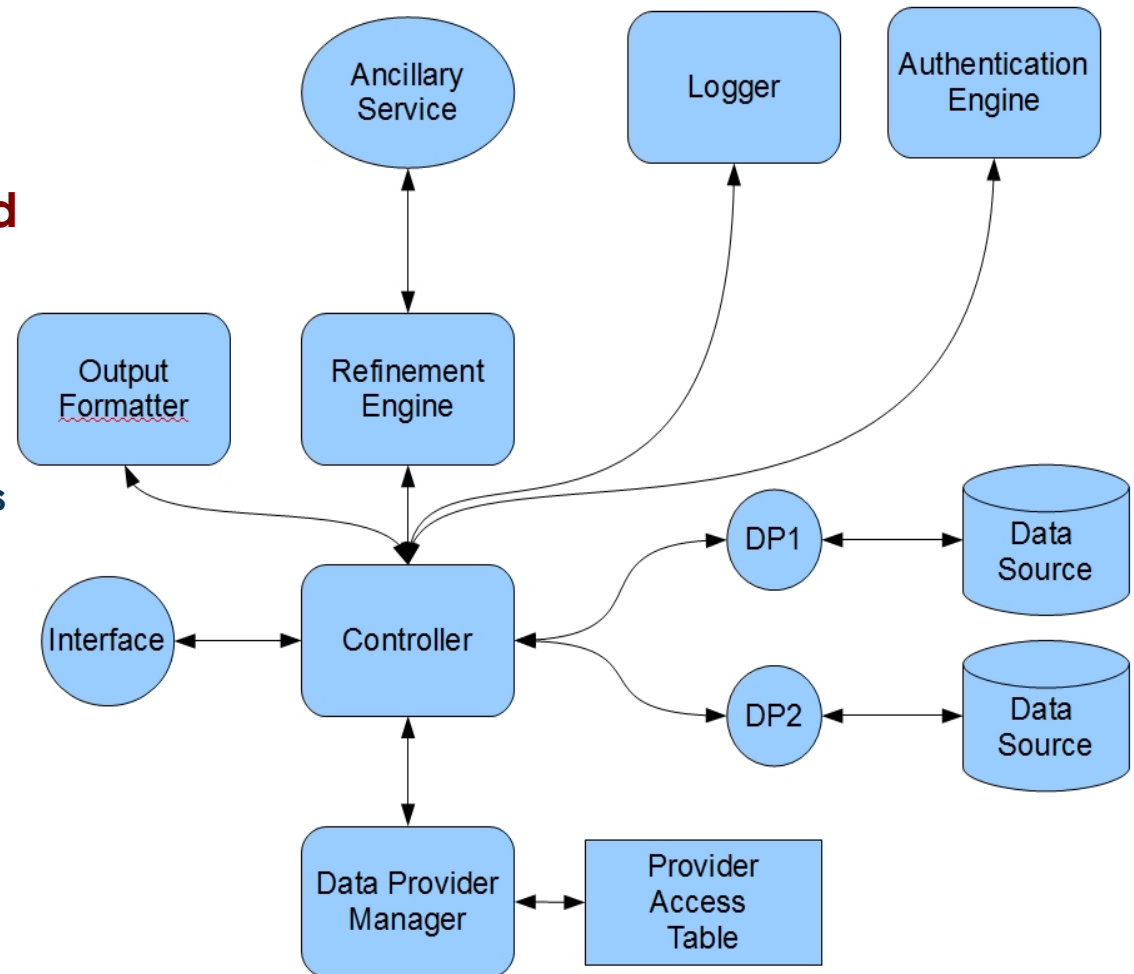
- **Metadata services:**
 - Heliophysics Event Catalogue (HEC)
 - Heliophysics Feature Catalogue (HFC)
 - Metadata Evaluation Service (MDES)
 - Context Service (CXs)
- **Discovery services:**
 - Instrument Capabilities Service (ICS)
 - Instrument Location Service (ILS)
 - Observation Coverage Service (OCS)
- **Data access services:**
 - Data Provider Access Service (DPAS)
- **Enabling services:**
 - Storage Service
 - Processing Service
 - Authentication/Authorization Service





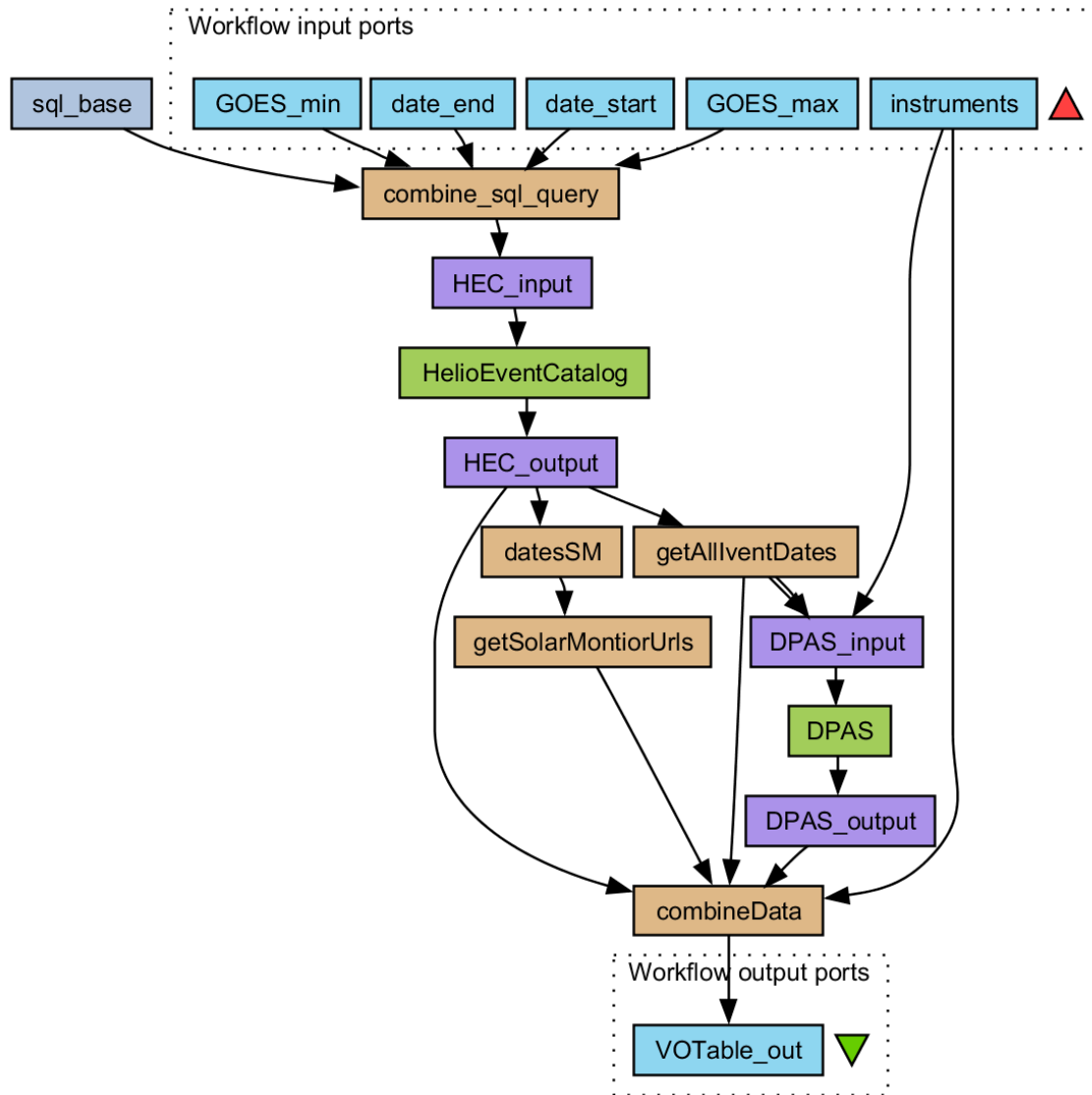
Centralized and Standardized Access to data sources:

- **Different sources**
 - Catalogues
 - Rich Interface data providers
- **Interaction with OCS**
 - To refine the queries
- **Redundant Data Sources**
 - For better reliability

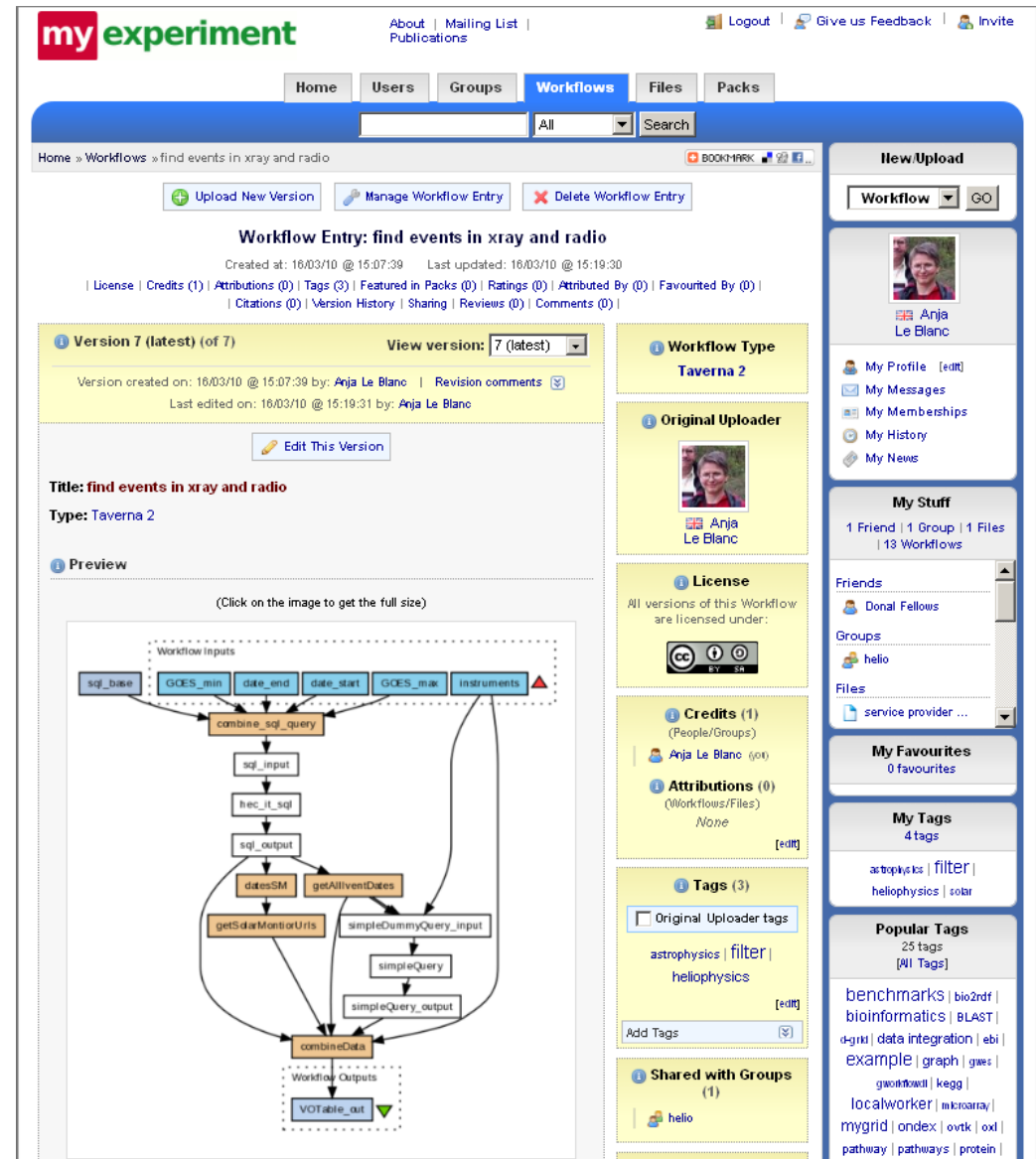




WORKFLOWS



- myExperiment exploits a very powerful feature of workflows – they can be shared and reused
- This can be combined with the more common features of social networking sites to build communities using and sharing common tools.
- myExperiment permits the sharing of workflows from different software, Kepler, Taverna, Triana etc...




The screenshot displays the myExperiment interface for a workflow entry titled "find events in xray and radio". The page includes a navigation bar with options like Home, Users, Groups, Workflows, Files, and Packs. The main content area shows the workflow's version history, metadata (created and last updated dates), and a detailed flowchart of the workflow process. The flowchart starts with "Workflow inputs" (sql_base, GOES_min, date_end, date_start, GOES_max, instruments) and proceeds through several steps: "combine_sql_query", "sql_input", "rec_it_sql", "sql_output", "datesSM", "getAllIventDates", "getStarMonitorUris", "simpleDummyQuery_input", "simpleQuery", "simpleQuery_output", and "combineData", finally leading to "Workflow Outputs" (VOTable_out). The right sidebar contains user profile information for Anja Le Blanc, including links to My Profile, My Messages, My Memberships, My History, and My News, as well as sections for My Stuff, Friends, Groups, Files, My Favourites, My Tags, and Popular Tags.



HELIO Advanced User Interface with jQuery



The screenshot displays the HELIO user interface. At the top, there is a dark blue header with the 'HELIO' logo in white. Below the header is a yellow navigation bar with four links: 'About Helio', 'How to use the site', 'Other Links', and 'Other Links'. The main content area features a vertical sidebar on the left with several icons: a clock, a satellite dish, a sun, a waveform, a calendar, a play button, and a button with the letter 'F'. The central part of the interface shows a search criteria dialog box titled 'Access Catalog'. This dialog box has a 'Top Five' section with a 'more' dropdown menu and five checkboxes labeled 'HEC', 'next important', 'next important', 'next important', and 'next important'. Below this section are two input fields for 'Add date range' and 'Add time range', each with 'From' and 'To' sub-fields. The dialog box also includes 'Add' and 'Cancel' buttons.



Time Range

Select Time Range

Top Five

From

To

Extend previous range


Subtract from previous range

Periodic Search

Repeat


From/Until

or Number of repetitions



Visual Search

Amount of data in current selection

From  To



Further elements



- Registry based on Astrogrid
- The Semantic VO
- Feature detection and tracking
- Education and Public Outreach



Contacts



<i>Institution</i>	<i>Representative</i>
University College London (MSSL)	Bob Bentley (PI)
Fachhochschule Nordwestschweiz (FHNW)	Andre Csillaghy
Observatoire de Paris (LESIA)	Jean Aboudarham
Universite Paul Sabatier Toulouse (CESR)	Christian Jacquy
Science and Technology Facilities Council (RAL)	Mike Hapgood
Universite Paris-Sud (IAS)	Karine Bocchialini
Istituto Nazionale di Astrofisica (Obs. Trieste)	Mauro Messerotti
University of Manchester	John Brooke
Trinity College Dublin	Brian Coghlan
Rensselaer Polytechnic Institute	Peter Fox
Lockheed Martin Space Systems Co. (LMATC)	Neal Hurlburt
NASA (<i>Heliophysics Science Division at GSFC</i>)	A. Roberts
European Space Agency (<i>Science Ops. Dept., Space Environment & Effects Dept.</i>)	L. Sanchez



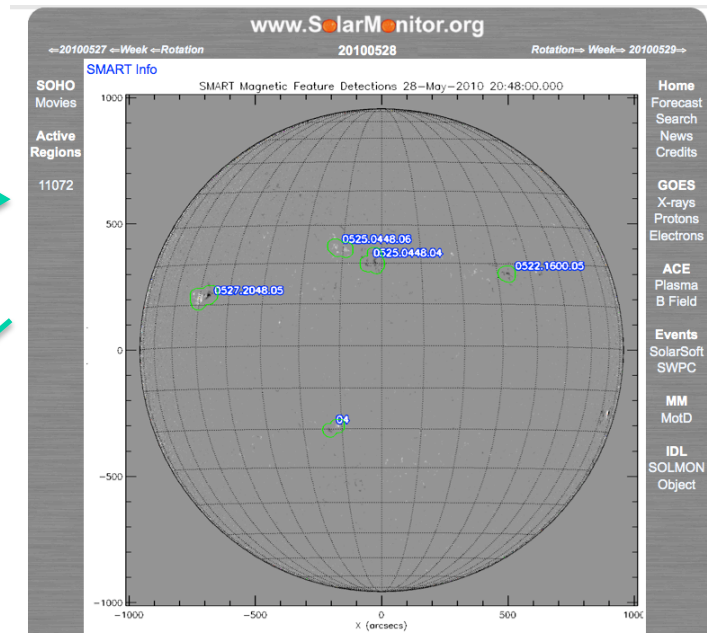
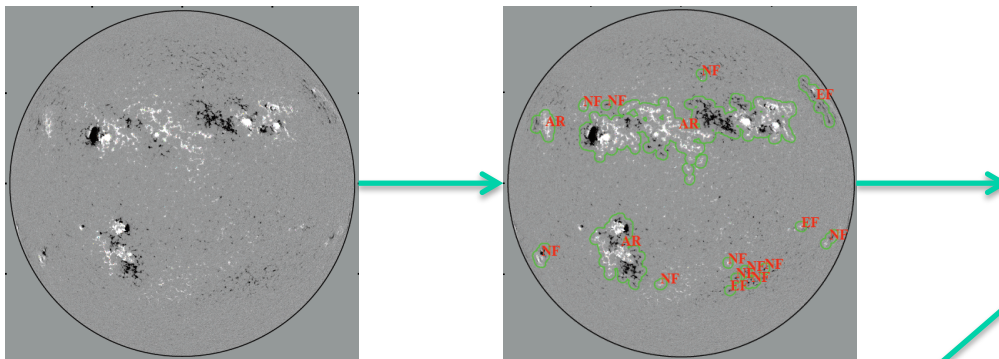
Semantic work - UCD & UType



We are building on previous work in VO to add semantics and models to the data. In this way the Semantic VO will be compatible with the wider VO movement

- **Semantic identifiers in VOTables**
- **Unified Content Descriptor (UCD) for high level mapping of parameters**
- **UCD is a controlled vocabulary list**
- **UType for mapping of parameter into a data model**
- **HELIO data model required**
- **Combination of UCD and UType provide a rich semantic basis**

SMART (SolarMonitor AR Tracking)



id	run_date	sc_arc_x	sc_arc_y	Feat_area	Feat_max_int
1	2010-06-09 21:16:51	-17.8166	-351.155	9.786	821.714
2	2010-06-09 21:16:51	-3.37461	-351.942	9.856	742.159
3	2010-06-09 21:16:52	9.97784	-352.144	9.8	736.317
4	2010-06-09 21:16:53	24.451	-353.546	10.22	694.35
5	2010-06-09 21:16:54	37.8423	-353.826	10.08	622.946
6	2010-06-09 21:16:55	51.0684	-352.984	9.618	689.682
7	2010-06-09 21:16:56	64.8875	-352.975	9.772	692.093
8	2010-06-09 21:16:57	78.5364	-353.843	9.912	798.685
9	2010-06-09 21:16:57	93.1801	-352.979	10.052	615.263
10	2010-06-09 21:16:58	106.303	-352.533	10.024	688.51
11	2010-06-09 21:16:59	120.964	-352.724	9.688	696.332
12	2010-06-09 21:17:00	134.88	-353.078	9.758	589.868
13	2010-06-09 21:17:01	149.258	-353.47	9.618	572.437
14	2010-06-09 21:17:02	162.91	-354.022	10.108	647.616
15	2010-06-09 21:17:02	178.563	-353.951	9.646	613.381
16	2010-06-09 21:17:03	192.388	-354.235	9.758	782.233
17	2010-06-09 21:17:04	207.215	-355.515	9.968	719.35
18	2010-06-09 21:17:05	221.242	-356.628	9.38	619.398
19	2010-06-09 21:17:06	234.051	-357.187	9.184	734.321

HELIO output

Reference: Higgins et al. (2010). Advances Space Research, in press.